GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 01:16:12 ; Search time 94 Seconds
                                            (without alignments)
                                            4761.303 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt..........tgaagaaaacggccccttga 1014

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                         %
Result   Query
    No.   Score   Match Length DB   ID   Description
    ----
        1   1014   100.0   1421   4   US-09-620-312D-111   Sequence 111, App
        2   1010.8   99.7   1344   3   US-09-190-965-2   Sequence 2, Appli
        3   1010.8   99.7   1344   4   US-09-470-253-2   Sequence 2, Appli
   C   4   51.6   5.1   7218   1   US-08-232-463-14   Sequence 14, Appl
        5   37.8   3.7   1457   4   US-09-214-307A-9   Sequence 9, Appli
   C   6   37.4   3.7   4103   4   US-09-620-312D-390   Sequence 390, App
        7   36.4   3.6   4533   3   US-08-726-214-5   Sequence 5, Appli
   C   8   35.6   3.5   577   4   US-09-513-057C-20   Sequence 20, Appl
   C   9   35.6   3.5   3707   3   US-09-276-531-42   Sequence 42, Appl
   C   10   35.6   3.5   5714   4   US-09-620-312D-393   Sequence 393, App
       11   35.2   3.5 12793   4   US-09-004-838-124   Sequence 124, App

	12	35	3.5	775	3	US-08-961-083-89	Sequence 89, Appl
	13	35	3.5	775	4	US-09-536-784-89	Sequence 89, Appl
	14	35	3.5	1696	3	US-08-961-083-217	Sequence 217, App
	15	35	3.5	1696	4	US-09-536-784-217	Sequence 217, App
	16	35	3.5	3763	1	US-07-792-865D-1	Sequence 1, Appli
C	17	35	3.5	13926	4	US-08-961-527-5	Sequence 5, Appli
C	18	34.8	3.4	1413	4	US-09-016-434-1382	Sequence 1382, Ap
	19	34.8	3.4	12734	4	US-09-344-456-1	Sequence 1, Appli
	20	34.8	3.4	14078	3	US-09-433-262-1	Sequence 1, Appli
	21	34.8	3.4	14078	4	US-09-702-330-1	Sequence 1, Appli
	22	34.8	3.4	14578	3	US-08-859-694-1	Sequence 1, Appli
C	23	34.6	3.4	64467	4	US-09-803-671B-3	Sequence 3, Appli
	24	34.4	3.4	22481	4	US-08-367-841A-43	Sequence 43, Appl
	25	34.4	3.4	22481	5	PCT-US95-07201-43	Sequence 43, Appl
	26	34.4	3.4	22484	4	US-09-875-223-2	Sequence 2, Appli
	27	34	3.4	3378	4	US-09-328-352-4107	Sequence 4107, Ap
C	28	34	3.4	43360	4	US-09-453-702B-206	Sequence 206, App
C	29	34	3.4	45325	4	US-09-453-702B-261	Sequence 261, App
	30	33.8	3.3	189	4	US-09-134-001C-1145	Sequence 1145, Ap
C	31	33.8	3.3	4233	3	US-09-056-105-27	Sequence 27, Appl
	32	33.6	3.3	1276	3	US-09-177-325-2	Sequence 2, Appli
	33	33.6	3.3	1276	3	US-09-411-812A-2	Sequence 2, Appli
	34	33.6	3.3	1276	4	US-09-590-113-2	Sequence 2, Appli
	35	33.6	3.3	1347	4	US-09-134-001C-1309	Sequence 1309, Ap
	36	33.4	3.3	1410	4	US-09-328-352-124	Sequence 124, App
C	37	33.4	3.3	1508	1	US-08-236-311-5	Sequence 5, Appli
C	38	33.4	3.3	1508	3	US-08-457-918-5	Sequence 5, Appli
	39	33.4	3.3	2643	4	US-09-486-072-6	Sequence 6, Appli
C	40	33.4	3.3	18627	4	US-08-961-527-113	Sequence 113, App
	41	33.2	3.3	378	4	US-09-252-991A-10313	Sequence 10313, A
C	42	33.2	3.3	813	4	US-09-252-991A-10547	Sequence 10547, A
	43	33.2	3.3	831	4	US-09-252-991A-10512	Sequence 10512, A
C	44	33.2	3.3	3253	4	US-09-333-214-4	Sequence 4, Appli
	45	33.2	3.3	8091	4	US-09-230-652-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-111
; Sequence 111, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing

```
;   APPLICANT:   Wang, Dunrui
;   APPLICANT:   Wang, Zhiwei
;   APPLICANT:   John Tillinghast
;   APPLICANT:   Drmanac, Radoje T.
;   TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;   TITLE OF INVENTION:  Polypeptides
;   FILE REFERENCE: 784CIP2B
;   CURRENT APPLICATION NUMBER: US/09/620,312D
;   CURRENT FILING DATE:  2000-07-19
;   PRIOR APPLICATION NUMBER: 09/552,317
;   PRIOR FILING DATE: 2000-04-25
;   PRIOR APPLICATION NUMBER: 09/488,725
;   PRIOR FILING DATE: 2000-01-21
;   NUMBER OF SEQ ID NOS: 1105
;   SOFTWARE: pt_FL_genes Version 1.0
;  SEQ ID NO 111
;   LENGTH: 1421
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (217)..(1230)
US-09-620-312D-111
```

```
   Query Match              100.0%;  Score 1014;  DB 4;  Length 1421;
   Best Local Similarity    100.0%;  Pred. No. 4.8e-292;
   Matches 1014;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276

Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336

Qy     121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396

Qy     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456

Qy     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516

Qy     301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576

Qy     361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636

Qy     421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
```

```
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       637  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC  696

Qy       481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  540
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       697  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  756

Qy       541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  600
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       757  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  816

Qy       601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  660
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       817  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  876

Qy       661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  720
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       877  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  936

Qy       721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  780
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       937  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  996

Qy       781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  840
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       997  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  1056

Qy       841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  900
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1057  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  1116

Qy       901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  960
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1117  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  1176

Qy       961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1014
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1177  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1230
```

RESULT 2
US-09-190-965-2
; Sequence 2, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
;   APPLICANT: Tang, Y. Tom
;   APPLICANT: Guegler, Karl J.
;   APPLICANT: Corley, Neil C.
;   APPLICANT: Gorgone, Gina A.
;   TITLE OF INVENTION: CALCIUM BINDING PROTEIN
;   FILE REFERENCE: PF-0635 US
;   CURRENT APPLICATION NUMBER: US/09/190,965
;   CURRENT FILING DATE: 1998-11-13
;   NUMBER OF SEQ ID NOS: 5
;   SOFTWARE: PERL Program
; SEQ ID NO 2
;   LENGTH: 1344

```
  Query Match              99.7%;   Score 1010.8;   DB 3;   Length 1344;
  Best Local Similarity    99.8%;   Pred. No. 4.2e-291;
  Matches 1012;   Conservative    0;   Mismatches    2;   Indels    0;   Gaps    0;

Qy          1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy         61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy        121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423

Qy        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

Qy        481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy        541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy        601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy        661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843
```

```
Qy    721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy    781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023

Qy    901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy    961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
```

RESULT 3
US-09-470-253-2
; Sequence 2, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 2
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 3734805
US-09-470-253-2

  Query Match   99.7%; Score 1010.8; DB 4; Length 1344;
  Best Local Similarity   99.8%; Pred. No. 4.2e-291;
  Matches 1012; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy      1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy     61 CTGAAAGACAATTTGGCCATTTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    184  CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy    121  GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244  GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy    181  CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    304  CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy    241  ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364  ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423

Qy    301  TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424  TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy    361  GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484  GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy    421  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

Qy    481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy    541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy    601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy    661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy    721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy    781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy    841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023

Qy    901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083
```

```
Qy      961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1084  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
```

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 1800 Diagonal Road, Suite 500
;   CITY: Alexandria
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/232,463
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/07/935,313
;   FILING DATE:
;   APPLICATION NUMBER: EP 91 114 300.6
;   FILING DATE: 26-AUG-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BENT, Stephen A.
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703)836-9300
;   TELEFAX: (703)683-4109
;   TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 7218 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   CLONE: pTZgpt-Fls
US-08-232-463-14

    Query Match   5.1%; Score 51.6; DB 1; Length 7218;
    Best Local Similarity 3.6%; Pred. No. 4.2e-05;
```

Matches   12;  Conservative  196;  Mismatches  130;  Indels   0;  Gaps   0;

Qy ·        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
              | || | | | |  |||| |   : ::: : : :::::   :: ::::: : :::::
Db      1456 AAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1397

Qy        61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
              ::::::: ::   ::  :    ::::::: :::: ::::::: ::: :::: :    ::::
Db      1396 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1337

Qy       121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
              :::: :   ::: :  : ::: :: ::::::::   : : :: : ::: ::::::::::::
Db      1336 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1277

Qy       181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
              :: ::::: :: ::   ::   :: ::: : :::::   : :: :: ::   :  :: :
Db      1276 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1217

Qy       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
              : :  :: :: :: ::   :: ::: :::   :::::::::::::::: : ::   ::: :
Db      1216 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1157

Qy       301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAG 338
              :: :: :   ::::::: ::: ::: :   ::::
Db      1156 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1119


RESULT 5
US-09-214-307A-9
; Sequence 9, Application US/09214307A
; Patent No. 6544516
; GENERAL INFORMATION:
;  APPLICANT:  NEUTEC PHARMA PLC
;  TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF INFECTIONS OF GRAM POSITIVE
;  TITLE OF INVENTION:  COCCI
;  FILE REFERENCE:  PM 259204
;  CURRENT APPLICATION NUMBER: US/09/214,307A
;  CURRENT FILING DATE:  1999-01-04
;  PRIOR APPLICATION NUMBER:  PCT/GB97/01830
;  PRIOR FILING DATE:  1997-07-07
;  PRIOR APPLICATION NUMBER:  GB9614274.0
;  PRIOR FILING DATE:  1996-07-06
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 1457
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-214-307A-9

  Query Match            3.7%;  Score 37.8;  DB 4;  Length 1457;
  Best Local Similarity  47.0%;  Pred. No. 0.23;
  Matches  150;  Conservative   0;  Mismatches  167;  Indels   2;  Gaps   1;

Qy       430 ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
              | |||| | | | | |||||| | ||||      |||| || ||||| ||||
Db       386 AATATGAGAACTGTAGTTGATCGACCTAGAACACAATATAAAAAAGTCGTCTTTAATAAT 445

```
Qy    490  CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC  549
           |||     | ||    |||| | |    | ||   | | |||| |  |    |  | ||
Db    446  TTATTTTATCAATTTAGTAAGGATGCCAACTTTGAACCTATTGCTTGTAGACCCTATCGT  505

Qy    550  TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA  609
           |  ||     | | | |   | |    | |||| | |   | | |    || | | |
Db    506  CCTCAAACAAAAGGGTCTGTTGAATCATTAGCTAAATTTGTTGAACAGCGTTTAAGACCA  565

Qy    610  CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT  669
           | |||| ||  |  | | || | || | |||| |||    ||| |||||||
Db    566  TACGATTATGAATTTTATGATGCTG--TAGAACTTATTGGGCTAGTAAACGATTTATGTC  623

Qy    670  ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC  729
           || || ||   ||| || | |  | | | | | ||| |    |  | |||| |
Db    624  ACGAATTGAATCACTTAGAAATTTCACAAGCAACAGAACAACGACCTATCGACGTTTTCA  683

Qy    730  ATCATGACAAAGTATATCA  748
           || |||| || | | ||
Db    684  ATTATGAAGAAAAGAACA  702
```

RESULT 6
US-09-620-312D-390/c
; Sequence 390, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 390
;   LENGTH: 4103

```
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (104)..(3493)
US-09-620-312D-390

   Query Match            3.7%;   Score 37.4;  DB 4;  Length 4103;
   Best Local Similarity  60.2%;  Pred. No. 0.53;
   Matches    62;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy          6 AAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAA 65
              | ||||| |||| | | | || |||||||||| | | | | ||||
Db       4091 ACAAATGAGAAAGTTTCATTTACCTCAAAAAAATCCAGGCTATACAAACAGACAACTGAA 4032

Qy         66 AGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC 108
              || || | || |||| |||| ||| | |||||| | ||
Db       4031 AGCCACATAGGAAATTTCCGAAACACAAAAGAAAAAGTCTCAC 3989


RESULT 7
US-08-726-214-5
; Sequence 5, Application US/08726214
; Patent No. 6107076
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Wei-Jen
;    APPLICANT:  Gilman, Alfred G.
;    TITLE OF INVENTION:  SOLUBLE MAMMALIAN ADENYLYL CYCLASE
;    TITLE OF INVENTION:  AND USES THEREFOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Arnold, White & Durkee
;     STREET:  P.O. Box 4433
;     CITY:  Houston
;     STATE:  Texas
;     COUNTRY:  United States of America
;     ZIP:  77210
;    COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/726,214
;     FILING DATE:  Concurrently Herewith
;     CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 60/005,498
;     FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;     NAME:  Highlander, Steven L.
;     REGISTRATION NUMBER:  37,642
;     REFERENCE/DOCKET NUMBER:  UTSD:450
;    TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (512) 418-3000
;     TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  5:
```

;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4533 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-726-214-5

Query Match   3.6%;   Score 36.4;   DB 3;   Length 4533;
Best Local Similarity   56.8%;   Pred. No. 1.1;
Matches   67; Conservative   0; Mismatches   51; Indels   0; Gaps   0;

```
Qy        718  CACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATG 777
               | ||  |    ||||||||    | |  ||||| |   || ||| ||||  ||||||| ||
Db       2644  CTCATCGCCACCATCATGCTGGTGCAGGTCAGCCACATGGTGAAGCTGACACTCATGCTG 2703

Qy        778  AACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGG 835
               | ||     | |      | | ||| || || || || ||     | || ||| ||| | |
Db       2704  CTCGTCACAGGCGCCGTGACTGCCATCAACCTGTATGCCTGGTGTCCTGTCTTTGATG 2761
```


RESULT 8
US-09-513-057C-20/c
; Sequence 20, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
;   APPLICANT: Wagner, et al.
;   TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND
PHOTOPERIODISM
;   FILE REFERENCE: 1505-54357
;   CURRENT APPLICATION NUMBER: US/09/513,057C
;   CURRENT FILING DATE: 2000-02-24
;   NUMBER OF SEQ ID NOS: 35
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 577
;   TYPE: DNA
;   ORGANISM: Lycopersicon esculentum
US-09-513-057C-20

Query Match   3.5%;   Score 35.6;   DB 4;   Length 577;
Best Local Similarity   51.2%;   Pred. No. 0.64;
Matches   83; Conservative   0; Mismatches   79; Indels   0; Gaps   0;

```
Qy        457  GAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTG 516
               || |||  | || ||| ||     |||| |  |  ||| |||||||||||  |       |
Db        223  GACCCAAATACCCAAAACACAATCTTTACATAGAAATCAAGAGATTTCTGAAGCACACAG 164

Qy        517  GAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACC 576
               ||     || || |   ||| |  | || || || || |||   | | |  | |
Db        163  AAGCAAAAAGATGTATAATTTCACAAAATTACTATTATATTTTTCTGTGATCATGTAAC 104

Qy        577  AGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
               || |     || | | ||||    ||||  | |||||
Db        103  AGGCCTTGTTGGTAAGCACAATAATATGAAGAAAGAGATTAC 62
```


RESULT 9

US-09-276-531-42/c
; Sequence 42, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
;   APPLICANT: Bandman, Olga
;   APPLICANT: Lal, Preeti
;   APPLICANT: Hillman, Jennifer L.
;   APPLICANT: Yue, Henry
;   APPLICANT: Reddy, Roopa
;   APPLICANT: Guegler, Karl J.
;   APPLICANT: Baughn, Mariah R.
;   TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
;   TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL
PROLIFERATION
;   NUMBER OF SEQUENCES: 134
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;   STREET: 3174 PORTER DRIVE
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/276,531
;   FILING DATE: Herewith
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/079,677
;   FILING DATE: March 27, 1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Lynn E. Murry, Ph.D.
;   REGISTRATION NUMBER: 42,918
;   REFERENCE/DOCKET NUMBER: PA-0008 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 42:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3707 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: CERVNOT01
;   CLONE: 936117
US-09-276-531-42

    Query Match   3.5%; Score 35.6; DB 3; Length 3707;
    Best Local Similarity   51.9%; Pred. No. 1.7;
    Matches   80; Conservative   0; Mismatches   74; Indels   0; Gaps   0;

```
Qy      445 TGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
            ||  ||| | |||     ||   | |||||    |||||||| || || |   | | ||
Db     3154 TGCTTTCAAAATGTGGAACAAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 3095

Qy      505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
            | || || |||| |     |||| | | | | | |        |    | ||| ||||
Db     3094 AATCAGCACTTGGATCTAATGACATATCTTTATAATACTTCCTCTGCAGATACATTCACT 3035

Qy      565 GATTTACTAACCAGACATAAAGTGTTGGTAGCAG 598
            | ||   | |   ||||| | ||| || |||
Db     3034 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 3001


RESULT 10
US-09-620-312D-393/c
; Sequence 393, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 393
;   LENGTH: 5714
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (272)..(4312)
US-09-620-312D-393

   Query Match            3.5%;  Score 35.6;  DB 4;  Length 5714;
   Best Local Similarity  51.9%;  Pred. No. 2.2;
```

```
Matches   80;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy       445 TGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
             ||  ||| | |||    ||  | |||||    |||||||| || || |   | | ||
Db      5233 TGCTTTCAAAATGTGGAACAAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 5174

Qy       505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
             ||| || |||| |    |||| | | | |    |        | |||  |||
Db      5173 AATCAGCACTTGGATCTAATGACATATCTTTGTAATACTTCCTCTGCAGATACATTCACT 5114

Qy       565 GATTTACTAACCAGACATAAAGTGTTGGTAGCAG 598
             | ||   | |   |||||  |  ||| || |||
Db      5113 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 5080


RESULT 11
US-09-004-838-124
; Sequence 124, Application US/09004838
; Patent No. 6350933
;  GENERAL INFORMATION:
;    APPLICANT:  Michelmore, Richard W.
;    APPLICANT:  Shen, Kathy
;    APPLICANT:  Meyers, Blake
;    TITLE OF INVENTION:  Procedures and Materials for
;    TITLE OF INVENTION:  Conferring Pest Resistance in Plants
;    NUMBER OF SEQUENCES:  140
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/004,838
;      FILING DATE:  09-JAN-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/781,734
;      FILING DATE:  10-JAN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Einhorn, Gregory P.
;      REGISTRATION NUMBER:  38,440
;      REFERENCE/DOCKET NUMBER:  023070-078810US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  124:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12793 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     FEATURE:
;        NAME/KEY:  -
;        LOCATION:  1..12793
;        OTHER INFORMATION:  /note= "RG2S"
US-09-004-838-124

  Query Match             3.5%;  Score 35.2;  DB 4;  Length 12793;
  Best Local Similarity  47.6%;  Pred. No. 4.4;
  Matches 101;  Conservative  10;  Mismatches  98;  Indels  3;  Gaps  1;

Qy       438 GAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAG 497
             ||||:| :  :    || |: ::|::::|     |   | || ||  |||      ||| ||
Db      5998 GAGARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAAAATGGAG 6057

Qy       498 AGATTTCTTTAAGTACGTGGAGTTGTCA---ACATTTGATATTGCTTCAGATGCCTTTGC 554
             || ||||| |    ||    || | |  |  | || ||||| ||| |  | |||
Db      6058 TTATATCTTTGATATTGTATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTAT 6117

Qy       555 TACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAA 614
             | |||| |  |||| |   | | | | |||||| ||      ||    ||| | ||||
Db      6118 AAGTTTTATTTATTTTGATATGAAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA 6177

Qy       615 TTACGACACTATTTTTGAAGACTATGAGAAAT 646
             | | ||| |||| ||| | | | | |||
Db      6178 ATCCAACAATATTAATCAAAAAGACCAMACAT 6209


RESULT 12
US-08-961-083-89
; Sequence 89, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;     APPLICANT:  Choi et. al.
;     TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;     NUMBER OF SEQUENCES:  452
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Human Genome Sciences, Inc.
;        STREET:  9410 Key West Avenue
;        CITY:  Rockville
;        STATE:  Maryland
;        COUNTRY:  USA
;        ZIP:  20850
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;        COMPUTER:  HP Vectra 486/33
;        OPERATING SYSTEM:  MSDOS version 6.2
;        SOFTWARE:  ASCII Text
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/08/961,083
;        FILING DATE:
;        CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:
;        FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
```

```
  Query Match              3.5%;  Score 35;  DB 3;  Length 775;
  Best Local Similarity   46.9%;  Pred. No. 1.1;
  Matches  143;  Conservative   0;  Mismatches  160;  Indels   2;  Gaps   1;

Qy       22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
            ||| || ||    | || ||||| ||     |||| |    ||| |    |  |   |
Db      263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322

Qy       82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
             |  ||    ||||| |||||| ||     |   ||||| |    |  || |  |    ||
Db      323 GCGCCAAGAGAAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382

Qy      142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCT 201
            | |  ||| |   |   |     | ||  | |||| | |  ||||  |   ||||
Db      383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442

Qy      202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
              |   ||   ||||  |  |  ||  | ||| || | |  ||  |     |||    ||
Db      443 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 500

Qy      262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
            ||| |  | |     | |||| |   || |||         ||| | | ||||  |
Db      501 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 560

Qy      322 CAGAT 326
            ||| |
Db      561 CAGGT 565
```

RESULT 13
US-09-536-784-89
; Sequence 89, Application US/09536784
; Patent No. 6573082
;   GENERAL INFORMATION:
;   APPLICANT: Choi et. al.
;   TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;   NUMBER OF SEQUENCES: 452
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA

```
  Query Match              3.5%;  Score 35;  DB 4;  Length 775;
  Best Local Similarity   46.9%;  Pred. No. 1.1;
  Matches  143;  Conservative    0;  Mismatches  160;  Indels    2;  Gaps     1;

Qy        22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
             ||| || ||   | || |||||| ||       ||| |    ||| |     | |   |
Db       263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322

Qy        82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
              |  ||    |||||  |||||  ||      |    ||||| |    | || | |   ||
Db       323 GCGCCAAGAGAAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382

Qy       142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCT 201
             | |  |||| |   | |   |      ||| |  | |||| |   | |||| |    ||||
Db       383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442

Qy       202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
               |   | |   |||| | | |  || | ||| || | | |  || |     ||| | ||
Db       443 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 500

Qy       262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
             ||| |  | |    | |||||| |  || |||    || ||| | | ||||  |
Db       501 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 560

Qy       322 CAGAT 326
             ||| |
Db       561 CAGGT 565
```

RESULT 14
US-08-961-083-217
; Sequence 217, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
;   APPLICANT: Choi et. al.
;   TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;   NUMBER OF SEQUENCES: 452
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;   COMPUTER: HP Vectra 486/33
;   OPERATING SYSTEM: MSDOS version 6.2
;   SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,083
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Brookes, A. Anders
;   REGISTRATION NUMBER: 36,373
;   REFERENCE/DOCKET NUMBER: PB340P2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (301) 309-8504
;   TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 217:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1696 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
US-08-961-083-217

    Query Match   3.5%; Score 35; DB 3; Length 1696;
    Best Local Similarity   46.9%; Pred. No. 1.7;
    Matches 143; Conservative   0; Mismatches 160; Indels   2; Gaps   1;

Qy   22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
             ||| || || | || ||||| || ||| | ||| | | | | |
Db   275 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 334

Qy   82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
             | || ||||| |||||| || | |||||| | | || | ||
Db   335 GCGCCAAGAGAAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394

Qy   142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCT 201
             | | |||| | | | | || | |||| | | |||| | ||||

```
Db      395 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454

Qy      202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
                |  | |   |||| | | |  ||  | ||| || | |  ||  |      |||  | ||
Db      455 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 512

Qy      262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
                ||| |  | |   |    ||||| |  || ||       |||| | | ||||  |
Db      513 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 572

Qy      322 CAGAT 326
                ||| |
Db      573 CAGGT 577


RESULT 15
US-09-536-784-217
; Sequence 217, Application US/09536784
; Patent No. 6573082
;   GENERAL INFORMATION:
;          APPLICANT: Choi et. al.
;          TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;          NUMBER OF SEQUENCES: 452
;          CORRESPONDENCE ADDRESS:
;               ADDRESSEE: Human Genome Sciences, Inc.
;               STREET: 9410 Key West Avenue
;               CITY: Rockville
;               STATE: Maryland
;               COUNTRY: USA
;               ZIP: 20850
;          COMPUTER READABLE FORM:
;               MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;               COMPUTER: HP Vectra 486/33
;               OPERATING SYSTEM: MSDOS version 6.2
;               SOFTWARE: ASCII Text
;          CURRENT APPLICATION DATA:
;               APPLICATION NUMBER: US/09/536,784
;               FILING DATE: 30-Oct-1997
;               CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;               APPLICATION NUMBER: 08/961,083
;               FILING DATE: OCT-30-1997
;          ATTORNEY/AGENT INFORMATION:
;               NAME: Michelle S. Marks
;               REGISTRATION NUMBER: 41,971
;               REFERENCE/DOCKET NUMBER: PB340P3
;          TELECOMMUNICATION INFORMATION:
;               TELEPHONE: (301) 309-8504
;               TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 217:
;          SEQUENCE CHARACTERISTICS:
;               LENGTH: 1696 base pairs
;               TYPE: nucleic acid
;               STRANDEDNESS: double
;               TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 217:
US-09-536-784-217
```

```
Query Match                3.5%;  Score 35;  DB 4;  Length 1696;
Best Local Similarity  46.9%;  Pred. No. 1.7;
Matches  143;  Conservative   0;  Mismatches  160;  Indels   2;  Gaps   1;

Qy     22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
          ||| || ||      | || |||||| ||        ||| |      ||| |    | |   |
Db    275 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 334

Qy     82 TTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
           |  ||    |||||| |||||| ||       |    |||||| |      | || | |   ||
Db    335 GCGCCAAGAGAAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394

Qy    142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCT 201
          | |  ||| |    |      | ||   | ||  | ||||  |  |  | ||||   |  ||||
Db    395 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454

Qy    202 CAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
           |   || |    |||| | | |  ||  | ||| || | |  ||  |      ||| |  ||
Db    455 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 512

Qy    262 CTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
          ||| |  | | |     | |||||  |   || |||        ||| |  | |||| |    |
Db    513 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 572

Qy    322 CAGAT 326
          ||| |
Db    573 CAGGT 577


Search completed: January  6, 2004, 03:19:48
Job time : 96 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 00:37:47 ; Search time 3965 Seconds
                                           (without alignments)
                                           10462.134 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt..........tgaagaaaacggccccttga 1014

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
                1: gb_ba:*
                2: gb_htg:*
                3: gb_in:*
                4: gb_om:*
                5: gb_ov:*
                6: gb_pat:*
                7: gb_ph:*
                8: gb_pl:*
                9: gb_pr:*
                10: gb_ro:*
                11: gb_sts:*
                12: gb_sy:*
                13: gb_un:*
                14: gb_vi:*
                15: em_ba:*
                16: em_fun:*
                17: em_hum:*
                18: em_in:*
                19: em_mu:*
                20: em_om:*
                21: em_or:*
                22: em_ov:*
                23: em_pat:*
                24: em_ph:*
                25: em_pl:*
                26: em_ro:*
                27: em_sts:*

```
28:   em_un:*
29:   em_vi:*
30:   em_htg_hum:*
31:   em_htg_inv:*
32:   em_htg_other:*
33:   em_htg_mus:*
34:   em_htg_pln:*
35:   em_htg_rod:*
36:   em_htg_mam:*
37:   em_htg_vrt:*
38:   em_sy:*
39:   em_htgo_hum:*
40:   em_htgo_mus:*
41:   em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1014	100.0	1014	6	AX105381	AX105381 Sequence
	2	1010.8	99.7	1344	6	AR097361	AR097361 Sequence
	3	1010.8	99.7	1344	6	AR203365	AR203365 Sequence
	4	1009.2	99.5	1491	9	BC010993	BC010993 Homo sapi
	5	992.8	97.9	2002	6	BD157871	BD157871 Primer fo
	6	992.8	97.9	2002	9	AK022639	AK022639 Homo sapi
	7	860.4	84.9	1359	10	BC016128	BC016128 Mus muscu
	8	858.8	84.7	1530	10	BC016546	BC016546 Mus muscu
	9	770.6	76.0	822	6	BD147463	BD147463 Primer fo
	10	684.6	67.5	831	6	BD079551	BD079551 Cancer-as
	11	582.6	57.5	1026	6	AX061831	AX061831 Sequence
	12	582.6	57.5	3281	6	AX082322	AX082322 Sequence
	13	582.6	57.5	3761	9	BC020570	BC020570 Homo sapi
	14	581.6	57.4	1680	9	AF151824	AF151824 Homo sapi
	15	581.6	57.4	3466	9	AF113536	AF113536 Homo sapi
	16	568.8	56.1	1947	10	BC020041	BC020041 Mus muscu
	17	567.2	55.9	2322	10	S51858	S51858 MO25 gene [
	18	552.8	54.5	3346	9	AK000804	AK000804 Homo sapi
	19	541.6	53.4	1053	6	AX105727	AX105727 Sequence
	20	533.6	52.6	851	9	AY211923	AY211923 Homo sapi
C	21	520.2	51.3	833	6	BD079552	BD079552 Cancer-as
	22	478.8	47.2	2366	5	BC044172	BC044172 Danio rer
	23	368.8	36.4	2991	10	BC029053	BC029053 Mus muscu
	24	365	36.0	158599	2	AC101941	AC101941 Mus muscu
	25	362.8	35.8	1894	3	AY128438	AY128438 Drosophil
	26	362.8	35.8	8324	2	AC019926	AC019926 Drosophil
C	27	362.8	35.8	201313	3	AC010688	AC010688 Drosophil
C	28	362.8	35.8	281101	3	AE003526	AE003526 Drosophil
	29	361.2	35.6	2213	3	AB000402	AB000402 Drosophil
	30	270	26.6	1452	9	AK026335	AK026335 Homo sapi
	31	255	25.1	1466	8	CPR238632	AJ238632 Chlorella
	32	240.8	23.7	3514	10	BC034159	BC034159 Mus muscu
	33	210.4	20.7	1474	8	AY088359	AY088359 Arabidops

```
         34   209.6   20.7    1570  8  AF428296          AF428296 Arabidops
         35   200.2   19.7    1032  8  AY055792          AY055792 Arabidops
         36   200.2   19.7    1519  8  AF380659          AF380659 Arabidops
    c    37    195    19.2     387  6  AX333971          AX333971 Sequence
    c    38    195    19.2     387  6  AX407834          AX407834 Sequence
    c    39   186.4   18.4   89137  8  AC006284          AC006284 Arabidops
    c    40   185.4   18.3   89807  9  AL138875          AL138875 Human DNA
    c    41   185.4   18.3  349980  6  AX711879          AX711879 Sequence
    c    42   185.4   18.3  349980  6  AX739961          AX739961 Sequence
    c    43   172.8   17.0  159863  9  AL136218          AL136218 Human DNA
         44   169.8   16.7     722  6  AX283846          AX283846 Sequence
         45   166.6   16.4     700  6  AX283847          AX283847 Sequence
```

                              ALIGNMENTS

RESULT 1
AX105381
LOCUS   AX105381   1014 bp   DNA   linear   PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0123552.
ACCESSION   AX105381
VERSION   AX105381.1 GI:13921508
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   den Daas,I. and Duecker,K.
  TITLE   Human paralogue of a head trauma induced cytoplasmatic calcium
            binding protein
  JOURNAL   Patent: WO 0123552-A 1 05-APR-2001;
            MERCK PATENT GmbH (DE)
FEATURES   Location/Qualifiers
     source   1. .1014
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS   1. .1014
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37735.1"
                     /db_xref="GI:13921509"
                     /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                     SLQAMKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                     NILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                     FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                     SENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                     FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                     TAP"
BASE COUNT   340 a   205 c   209 g   260 t
ORIGIN

     Query Match   100.0%; Score 1014; DB 6; Length 1014;
     Best Local Similarity 100.0%; Pred. No. 6.9e-238;
     Matches 1014; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
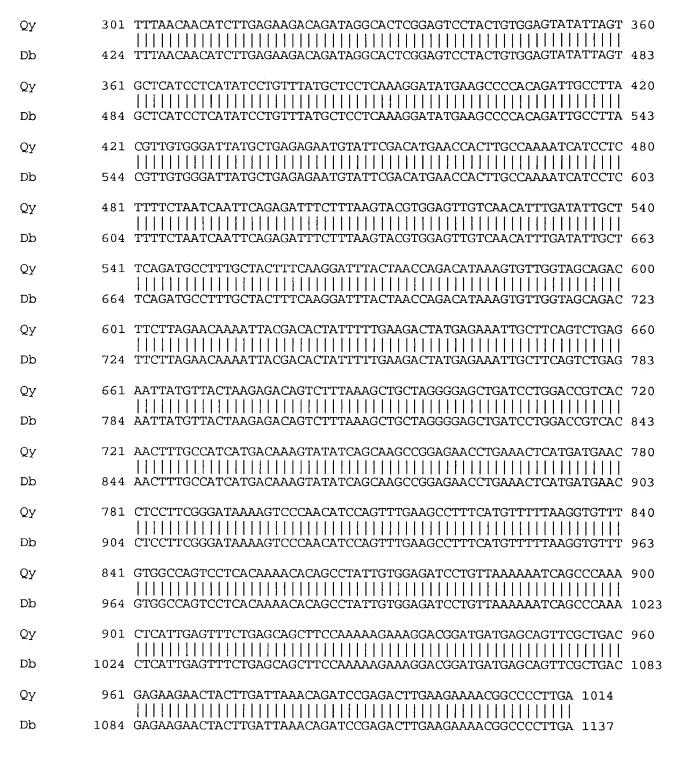Qy     1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

Qy    61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240

Qy   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300

Qy   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360

Qy   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420

Qy   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480

Qy   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540

Qy   541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600

Qy   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660

Qy   661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720

Qy   721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780

Qy   781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
```

```
Qy      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900

Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
```

RESULT 2
AR097361

```
LOCUS       AR097361                 1344 bp    DNA        linear    PAT 14-FEB-2001
DEFINITION  Sequence 2 from patent US 6071721.
ACCESSION   AR097361
VERSION     AR097361.1  GI:12806091
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1344)
  AUTHORS   Tang,Y.Tom., Guegler,K.J., Corley,N.C. and Gorgone,G.A.
  TITLE     Calcium binding protein
  JOURNAL   Patent: US 6071721-A 2 06-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="unknown"
BASE COUNT        450 a      261 c      280 g      353 t
ORIGIN
```

  Query Match   99.7%; Score 1010.8; DB 6; Length 1344;
  Best Local Similarity   99.8%; Pred. No. 4.2e-237;
  Matches 1012; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy       61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            |||||  ||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423
```

```
Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

Qy      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy      661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy      721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy      781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023

Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
```

RESULT 3
AR203365
LOCUS   AR203365   1344 bp   DNA   linear   PAT 20-JUN-2002
DEFINITION Sequence 2 from patent US 6365371.
ACCESSION   AR203365
VERSION   AR203365.1 GI:21499736
KEYWORDS   .
SOURCE   Unknown.

```
ORGANISM   Unknown.
           Unclassified.
REFERENCE  1  (bases 1 to 1344)
   AUTHORS  Tang,Y.Tom., Guegler,K.J., Corley,N.C. and Gorgone,G.A.
   TITLE    Calcium binding protein
   JOURNAL  Patent: US 6365371-A 2 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="unknown"
BASE COUNT     450 a    261 c    280 g    353 t
ORIGIN
```

```
  Query Match              99.7%;  Score 1010.8;  DB 6;  Length 1344;
  Best Local Similarity    99.8%;  Pred. No. 4.2e-237;
  Matches 1012;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy       61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db      304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

Qy      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724   TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy    661   AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784   AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy    721   AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844   AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy    781   CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904   CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy    841   GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964   GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023

Qy    901   CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1024  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy    961   GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1084  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
```

RESULT 4
BC010993
LOCUS   BC010993   1491 bp   mRNA   linear   PRI 25-JUL-2001
DEFINITION   Homo sapiens, hypothetical protein FLJ12577, clone MGC:15031
             IMAGE:3956127, mRNA, complete cds.
ACCESSION   BC010993
VERSION   BC010993.1 GI:15012172
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 1491)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-JUL-2001) National Institutes of Health, Mammalian
             Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Institute for Systems Biology
             http://www.systemsbiology.org
             contact: amadan@systemsbiology.org
             Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia

Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 25 Row: k Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10434146.

```
FEATURES             Location/Qualifiers
     source          1. .1491
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="LocusID:81617"
                     /db_xref="taxon:9606"
                     /clone="MGC:15031 IMAGE:3956127"
                     /tissue_type="Placenta, choriocarcinoma"
                     /clone_lib="NIH_MGC_21"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             416. .1285
                     /codon_start=1
                     /product="hypothetical protein FLJ12577"
                     /protein_id="AAH10993.1"
                     /db_xref="GI:15012173"
                     /translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
                     EKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
                     RHEPLAKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT
                     IFEDYEKLLQSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
                     PNIQFEAFHVFKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNY
                     LIKQIRDLKKTAP"
BASE COUNT      503 a      290 c      305 g      393 t
ORIGIN
```

  Query Match   99.5%;   Score 1009.2; DB 9; Length 1491;
  Best Local Similarity   99.7%;   Pred. No. 1e-236;
  Matches 1011; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      272 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 331

Qy       61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      332 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 391

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      392 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAGAA 451

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      452 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 511

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db      512 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGAAAAAAAAGATGTGACCCAGATA 571

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
```

```
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     572  TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 631

Qy     361  GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     632  GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 691

Qy     421  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     692  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 751

Qy     481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     752  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 811

Qy     541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     812  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 871

Qy     601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     872  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 931

Qy     661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     932  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 991

Qy     721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     992  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 1051

Qy     781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1052  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1111

Qy     841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1112  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1171

Qy     901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1172  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1231

Qy     961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1232  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1285
```

RESULT 5
BD157871

LOCUS   BD157871   2002 bp   DNA   linear   PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD157871
VERSION   BD157871.1 GI:27863629
KEYWORDS   JP 2002191363-A/12714.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens

    Query Match   97.9%; Score 992.8; DB 6; Length 2002;
    Best Local Similarity   99.8%; Pred. No. 1.1e-232;
    Matches 994; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

Qy   79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

Qy   139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180

Qy   199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

Qy   259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

Qy   319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db      301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

Qy      379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

Qy      439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480

Qy      499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

Qy      559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600

Qy      619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy      679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy      739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780

Qy      799 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 858
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 840

Qy      859 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 900

Qy      919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960

Qy      979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            |||||||||||||||||||||||||||||||||||||
Db      961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
```

RESULT 6
AK022639
LOCUS   AK022639   2002 bp   mRNA   linear   PRI 01-AUG-2002
DEFINITION   Homo sapiens cDNA FLJ12577 fis, clone NT2RM4001047, highly similar
             to MO25 PROTEIN.
ACCESSION   AK022639
VERSION   AK022639.1 GI:10434146
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens

```
                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1
   AUTHORS      Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
                Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
                Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
                Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
                Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
                Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
                Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
   TITLE        NEDO human cDNA sequencing project
   JOURNAL      Unpublished
REFERENCE       2  (bases 1 to 2002)
   AUTHORS      Isogai,T. and Otsuki,T.
   TITLE        Direct Submission
   JOURNAL      Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
                Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
                (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT         NEDO human cDNA sequencing project supported by Ministry of
                International Trade and Industry of Japan; cDNA full insert
                sequencing: Research Association for Biotechnology; cDNA library
                construction, 5'- & 3'-end one pass sequencing and clone selection:
                Helix Research Institute (supported by Japan Key Technology Center
                etc.) and Department of Virology, Institute of Medical Science,
                University of Tokyo.
FEATURES             Location/Qualifiers
     source          1..2002
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NT2RM4001047"
                     /cell_line="NT2"
                     /cell_type="teratocarcinoma"
                     /clone_lib="NT2RM4"
                     /note="cloning vector: pME18SFL3~mRNA from uninduced NT2
                     neuronal precursor cells."
     CDS             127..996
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB14147.1"
                     /db_xref="GI:10434147"
                     /translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
                     GKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
                     RHEPLVKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT
                     IFEDYEKLLQSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
                     PNIQFEAFHVFKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNY
                     LIKQIRDLKKTAP"
BASE COUNT        594 a     418 c     463 g     527 t
ORIGIN

   Query Match                97.9%;  Score 992.8;  DB 9;  Length 2002;
   Best Local Similarity      99.8%;  Pred. No. 1.1e-232;
   Matches  994;  Conservative   0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
```

```
Qy     79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

Qy    139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180

Qy    199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
          ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

Qy    259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

Qy    319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

Qy    379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

Qy    439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480

Qy    499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

Qy    559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600

Qy    619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy    679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy    739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780

Qy    799 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 858
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 840

Qy    859 ACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 900
```

```
Qy        919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960

Qy        979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              ||||||||||||||||||||||||||||||||||||
Db        961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
```

RESULT 7
BC016128

```
LOCUS       BC016128                1359 bp    mRNA    linear   ROD 16-APR-2003
DEFINITION  Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:28889
            IMAGE:4911640), complete cds.
ACCESSION   BC016128
VERSION     BC016128.1  GI:16359341
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1359)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE   22388257
   PUBMED    12477932
REFERENCE   2  (bases 1 to 1359)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
```

FEATURES   Location/Qualifiers
     source   1. .1359
                        /organism="Mus musculus"
                        /mol_type="mRNA"
                        /strain="FVB/N"
                        /db_xref="taxon:10090"
                        /clone="MGC:28889 IMAGE:4911640"
                        /tissue_type="Salivary gland, 10 week old female mouse"
                        /clone_lib="NCI_CGAP_SG2"
                        /lab_host="DH10B"
                        /note="Vector: pCMV-SPORT6"
     gene   1. .1359
                        /gene="1500031K13Rik"
                        /note="synonyms: 4930520C08Rik, 2810425O13Rik"
                        /db_xref="LocusID:69008"
                        /db_xref="MGI:1916258"
     CDS   262. .1266
                        /codon_start=1
                        /product="1500031K13Rik protein"
                        /protein_id="AAH16128.1"
                        /db_xref="GI:16359342"
                        /db_xref="LocusID:69008"
                        /translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
                        AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
                        RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
                        QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
                        YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
                        FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
                        "
BASE COUNT   418 a   301 c   294 g   346 t
ORIGIN

  Query Match   84.9%; Score 860.4; DB 10; Length 1359;
  Best Local Similarity   90.5%; Pred. No. 3.2e-200;
  Matches 918; Conservative   0; Mismatches   96; Indels   0; Gaps   0;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||

```
Db    253 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 312

Qy     61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 372

Qy    121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
          || ||||| ||||| |||||||||||||||| ||||||||||||||| || |||||| ||
Db    373 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 432

Qy    181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          ||||| ||||||||||||||||||||| || || || ||||||||||| || |||| |||
Db    433 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 492

Qy    241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
          ||||| |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    493 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 552

Qy    301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          || |||||||||| |||||||||||| ||  || |||||||||||||| |||||  | ||
Db    553 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 612

Qy    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          |||||||||| ||||||||||||||| ||||||| |||||||||||||||||||||||||
Db    613 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 672

Qy    421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
          || |||||||||||||| ||||| |||||||||||||| |||||||||||||||||||| 
Db    673 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 732

Qy    481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          ||||||||||| |||||||||||||| ||||| || ||| |||| || ||||||||| ||
Db    733 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 792

Qy    541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
          |||||||||| |||||||||| |||||||| ||||||||||||||||| ||||||||||
Db    793 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 852

Qy    601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          ||||||||||||||||| ||||||||||||||||||||||||||| |||| || ||||||
Db    853 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 912

Qy    661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
          || |||||| || ||||||||| |||||| ||||||| |||||||||||||||||||  ||
Db    913 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 972

Qy    721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
          || ||  ||||| ||||||||||||||||||||||| ||||||||||||||||||||||
Db    973 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1032

Qy    781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
          || |||||| || |||||||||||||||| || |||||||| ||||| ||||||||||||
Db   1033 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1092

Qy    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
          ||||||||| || ||||||||  |||||| |||||||| |||||||||||||||||||||
Db   1093 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1152
```

```
Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
            |||||||||||||||||||||||||| || |||||||||| || || |||||||| ||||||
Db     1153 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1212

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            ||||||||||| |||||||||||| ||||||||||||||||| | ||||| |||
Db     1213 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1266
```

RESULT 8
BC016546

```
LOCUS       BC016546                1530 bp    mRNA     linear   ROD 16-APR-2003
DEFINITION  Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:27972
            IMAGE:3595339), complete cds.
ACCESSION   BC016546
VERSION     BC016546.1  GI:16741456
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1530)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1530)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 35 Row: m Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

```
FEATURES             Location/Qualifiers
     source          1. .1530
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:27972 IMAGE:3595339"
                     /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
                     ductal carcinoma. 5 month old virgin mouse."
                     /clone_lib="NCI_CGAP_Mam6"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1530
                     /gene="1500031K13Rik"
                     /note="synonyms: 4930520C08Rik, 2810425O13Rik"
                     /db_xref="LocusID:69008"
                     /db_xref="MGI:1916258"
     CDS             279. .1283
                     /codon_start=1
                     /product="1500031K13Rik protein"
                     /protein_id="AAH16546.1"
                     /db_xref="GI:16741457"
                     /db_xref="LocusID:69008"
                     /translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
                     AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
                     RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
                     QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
                     YVTKRQSLKLRGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
                     FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
                     "
BASE COUNT      498 a    313 c    326 g     393 t
ORIGIN
```

Query Match   84.7%;   Score 858.8; DB 10;   Length 1530;
Best Local Similarity   90.4%;   Pred. No. 8e-200;
Matches 917; Conservative   0; Mismatches   97; Indels   0; Gaps   0;

```
Qy         1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       270 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 329
```

```
Qy     61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
          |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db    330 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 389

Qy    121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
          || ||||||  |||||  ||||||||||||||||| |||||||||||  || ||||| || |||
Db    390 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 449

Qy    181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          |||||  ||||||||||||||||||||  || || || |||||||||||  ||  |||| |||
Db    450 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 509

Qy    241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
          |||||  |||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db    510 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 569

Qy    301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          || ||||||||| |||||||||||||| || || ||| ||||||||| ||||| || ||| |||
Db    570 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 629

Qy    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          ||||||||| |||||||||||||||||  |||||| ||||||||||||||||||||||||||
Db    630 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 689

Qy    421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
          || ||||||||||||||| ||||| |||||||||||||| |||||||||||||||||||||
Db    690 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 749

Qy    481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          |||||||||| |||||||||||||||| ||||| || ||| |||| || |||||||| |||
Db    750 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 809

Qy    541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
          ||||||||||| |||||||| |||||||| ||||||||  ||||||||||| |||||||||
Db    810 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 869

Qy    601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          |||||||||||||||| |||||||||||||||||||||||||||||  || || ||||||
Db    870 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 929

Qy    661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
          || |||||| || ||||||||| ||||||||||||| ||| ||| ||||||||||||| |||
Db    930 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCGAGGTGAGCTGATCCTGGACCGCCAC 989

Qy    721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
          || ||  |||| |||||||||||||||||||||||| |||||||||||||||||||||||
Db    990 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1049

Qy    781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
          || |||||  || |||||||||||||||||  || ||||||| ||||| ||||||||||||
Db   1050 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1109

Qy    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
          |||||||||  |  ||||||| |||||||| |||||||| ||||||||||||||||||||||
Db   1110 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1169
```

```
Qy        901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
              ||||||||||||||||||||||||| || ||||||||||| || || ||||||||| ||||||
Db       1170 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1229

Qy        961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              ||||||||||| |||||||||||||| |||||||||||||||| | |||||| |||
Db       1230 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1283
```

RESULT 9
BD147463

```
LOCUS       BD147463                822 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD147463
VERSION     BD147463.1  GI:27853221
KEYWORDS    JP 2002191363-A/2306.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 2306 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/2306
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
            10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                      Location/Qualifiers
            FT   source          1. .822
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      268 a     164 c     171 g     216 t       3 others
ORIGIN
```

```
  Query Match              76.0%;  Score 770.6;  DB 6;  Length 822;
  Best Local Similarity    98.5%;  Pred. No. 3.4e-178;
  Matches  798;  Conservative   0;  Mismatches   10;  Indels   2;  Gaps   2;
```

```
Qy         19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
```

```
Qy    79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

Qy   139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180

Qy   199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
         ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

Qy   259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

Qy   319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

Qy   379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

Qy   439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480

Qy   499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

Qy   559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600

Qy   619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy   679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy   739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
         |||||||||||||||||||||||||||| |||||||||||||| ||||||||| |||||
Db   721 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778

Qy   799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
         |||||||||||||||||||  |  |  | ||
Db   779 CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808
```

RESULT 10
BD079551
LOCUS   BD079551   831 bp   DNA   linear   PAT 27-AUG-2002

DEFINITION   Cancer-associated nucleic acids and polypeptides.
ACCESSION   BD079551
VERSION   BD079551.1 GI:22625154
KEYWORDS   JP 2001516009-A/217.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 831)
  AUTHORS   Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I.,
             OGhare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.
  TITLE   Cancer-associated nucleic acids and polypeptides
  JOURNAL   Patent: JP 2001516009-A 217 25-SEP-2001;
             LUDWIG INSTITUTE FOR CANCER RESEARCH
COMMENT   OS   Homo sapiens (human)
             PN   JP 2001516009-A/217
             PD   25-SEP-2001
             PF   15-JUL-1998 JP 2000503425
             PR   17-JUL-1997 US   08/896164,10-OCT-1997 US   60/061599 PR
             10-OCT-1997 US   60/061765,10-OCT-1997 US   08/948705 PR
             11-OCT-1997 GB   9721697.2,22-JUN-1998 US   09/102322 PI   LLOYD
             J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO   PI   TSENG
             CHEN,
             PI   IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI
               OZLEM TURECI,
             PI   UGUR SAHIN
             PC
             G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC
              A61P35/00,
             PC   C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
             PC   C12N15/00
             CC   Cancer-associated nucleic acids and polypeptides. FH   Key
                     Location/Qualifiers
             FT   source   1. .831
             FT   /organism='Homo sapiens (human)'.
FEATURES   Location/Qualifiers
     source   1. .831
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT   285 a   165 c   167 g   209 t   5 others
ORIGIN

  Query Match   67.5%; Score 684.6; DB 6; Length 831;
  Best Local Similarity 96.1%; Pred. No. 4.1e-157;
  Matches 764; Conservative   0; Mismatches   23; Indels   8; Gaps   6;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   37 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 156

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    157 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 216

Qy    181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 276

Qy    241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    277 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 336

Qy    301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    337 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 396

Qy    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    397 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456

Qy    421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 516

Qy    481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    517 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 576

Qy    541 TCAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTTGGTAGC-AG 598
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||| ||
Db    577 TCAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTTGGTAGCAAG 636

Qy    599 ACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 658
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    637 ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 696

Qy    659 AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 716
          || |||||||||| ||||||||||| |||||| ||||| |||| ||||| ||||||||
Db    697 AGAAATTATGTTACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAACTGATTCTGGACCG 756

Qy    717 TCACAACTTTGCCATC-ATGACAAAGTATATCAGCAAGCC---GGAGAACCTGAAACTCA 772
          ||| ||||||||||| | | |||||| |||||| |||||| ||| | | || |
Db    757 TCANAACTTTGCCATCAANGCAAAGTTTATCAACAAGCCNGGGGAAACCGGAAACNCAA 816

Qy    773 TGATGAACCTCCTTC 787
          || |||||||||||
Db    817 GGAGGAACCTCCTTC 831
```

RESULT 11
AX061831
LOCUS   AX061831   1026 bp   DNA   linear   PAT 24-JAN-2001
DEFINITION Sequence 1 from Patent WO0078947.
ACCESSION   AX061831
VERSION   AX061831.1 GI:12539911
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

FEATURES   Location/Qualifiers
     source   1. .1026
                         /organism="Homo sapiens"
                         /mol_type="genomic DNA"
                         /db_xref="taxon:9606"
     CDS   1. .1026
                         /note="unnamed protein product"
                         /codon_start=1
                         /protein_id="CAC25030.1"
                         /db_xref="GI:12539912"
                         /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                         KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                         NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                         LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                         HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                         VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                         RPAQQEA"
BASE COUNT   359 a   199 c   203 g   265 t
ORIGIN

   Query Match   57.5%;   Score 582.6;   DB 6;   Length 1026;
   Best Local Similarity   74.7%;   Pred. No. 4.3e-132;
   Matches 748; Conservative   0; Mismatches 244;   Indels   9; Gaps   1;

Qy   18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
             |||| | || || |||||| |||||||||| ||||||| | ||||| || | ||||
Db   12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 71

Qy   78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
             || ||||||||||||||| ||||| |||| ||||||| ||||||||||| ||
Db   72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131

Qy   129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
             ||| ||| || |||||||||||||||| ||| ||||| || ||||| || | ||
Db   132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191

Qy   189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
             |||||||||| ||||| || || |||||||||||| | ||||| || || || ||| |
Db   192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 251

Qy   249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
             |||||| | |||||| || |||||||||| ||||||||| ||| | || || || |||||
Db   252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 311

Qy   309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
             || | |||||||| || || ||| | |||||||| || || || | | |||| | |
Db   312 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 371

Qy   369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
             |||| |||| ||| | | ||||| |||||| | ||| | || || || |||||||
Db   372 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 431

```
Qy    429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
          || ||| | |||||||| || ||||||||||||||||| ||||||| | | || |
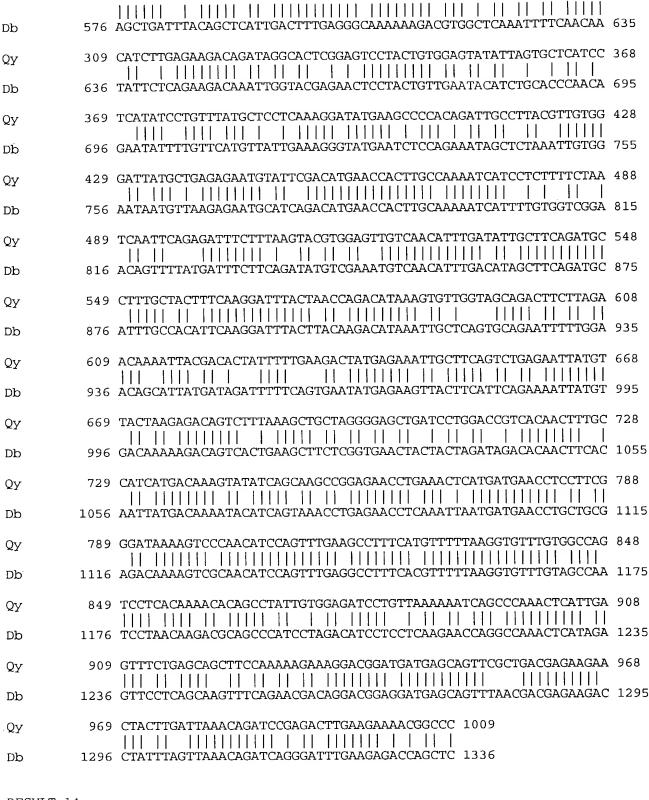Db    432 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAATCATTTTGTGGTCGGA 491


Qy    489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
          || || ||||||||| | || || || ||||||||||||||| || ||||||||||||
Db    492 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 551


Qy    549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
          ||||| || ||||||||||||| || |||||||||| | | ||||| || || ||
Db    552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 611


Qy    609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
          ||| |||| || | |||| || ||||||||| || ||||| || || ||||||||
Db    612 ACAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 671


Qy    669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
          || || ||||||| | ||||| || || || || | || || | |||||||| |
Db    672 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 731


Qy    729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
          || ||||||||| || ||||| || || ||||||||| ||| | |||||||||| || ||
Db    732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 791


Qy    789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
          || ||||||| ||||||||||||||||| |||||||| |||||||||||||||| ||||
Db    792 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 851


Qy    849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
          |||| |||| || ||||| || | || |||||| | || || ||| |||||||||| ||
Db    852 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 911


Qy    909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
          ||| || |||| || || || || || |||||||| |||||||||   |||||||||
Db    912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 971


Qy    969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
          ||| || ||||||||||| | || ||||||| | ||| |
Db    972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012
```

RESULT 12
AX082322
LOCUS   AX082322   3281 bp   DNA   linear   PAT 28-FEB-2001
DEFINITION Sequence 26 from Patent WO0111032.
ACCESSION   AX082322
VERSION   AX082322.1 GI:13184499
KEYWORDS   .
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hodgson,D.M., Lincoln,S.E., Russo,F.D., Spiro,P.A., Banville,S.C.,
            Bratcher,S.R., Dufour,G.E., Cohen,H.J., Rosen,B.H., Chalup,M.S.,
            Hillman,J.L., Jones,A.L., Yu,J.Y., Greenawalt,L.B., Panzer,S.R.,

  Query Match   57.5%;   Score 582.6; DB 6;   Length 3281;
  Best Local Similarity   74.7%;   Pred. No. 4.3e-132;
  Matches 748; Conservative   0; Mismatches 244; Indels   9; Gaps   1;

Qy   18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
            |||| | || || ||||||| |||||||||| |||||||| | ||||| || | ||||
Db   101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 160

Qy   78 CATTTTGGAAAAGCAAGAC---------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
            || ||||||||||||||| ||||| |||| |||||| ||||||||||| ||
Db   161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 220

Qy   129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
            ||| ||| || ||||||||||||||||| ||| |||||| || ||||| || | ||
Db   221 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 280

Qy   189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
            ||||||||| ||||| || || ||||||||||| | ||||| || || || ||| |
Db   281 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 340

Qy   249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
            |||||| | |||||| ||| |||||||||||||| ||| | || || || |||||
Db   341 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 400

Qy   309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
            || | ||||||||| || || || | | ||||||||| || || || | | ||| |
Db   401 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 460

Qy   369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
            |||| |||| ||| | | ||||| |||||| ||| | | || | || | |||||
Db   461 GAATATTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 520

Qy   429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
            || ||| | |||||||||| || ||||||||||||||||||||||||||||·| | || |
Db   521 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 580

Qy   489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
            || || |||||||| | || || || || ||||||||||||||| || |||||||||||
Db   581 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 640

Qy   549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
            |||||| || |||||||||||||| || |||||||||| || ||||| || || ||
Db   641 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 700

```
Qy    609  ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
           |||  |||| || |    ||||    || |||||||| || |||||| || || ||||||||
Db    701  ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 760

Qy    669  TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
           || || |||||||||  | ||||| || || || ||  | || ||  | |||||||| |
Db    761  GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 820

Qy    729  CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
           || |||||||||| || |||||| || || |||||||| |||  | |||||||||| || ||
Db    821  AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 880

Qy    789  GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
           || ||||||||| |||||||||||||||| ||||||||| ||||||||||||||||| ||||
Db    881  AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 940

Qy    849  TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
           |||| |||| || ||||| || | || |||||  | || || ||| |||||||||| ||
Db    941  TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1000

Qy    909  GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
           ||| || ||||  || || || || || |||||||| ||||||||||      ||||||||||
Db   1001  GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1060

Qy    969  CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
           ||| ||  |||||||||||| | || |||||||| | | || |
Db   1061  CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1101
```

RESULT 13
BC020570
LOCUS   BC020570   3761 bp   mRNA   linear   PRI 22-JAN-2002
DEFINITION Homo sapiens, MO25 protein, clone MGC:21631 IMAGE:4397573, mRNA,
            complete cds.
ACCESSION   BC020570
VERSION   BC020570.1 GI:18088260
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3761)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www-shgc.stanford.edu

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 27 Row: d Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction.

```
FEATURES             Location/Qualifiers
     source          1. .3761
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="LocusID:51719"
                     /db_xref="taxon:9606"
                     /clone="MGC:21631 IMAGE:4397573"
                     /tissue_type="Duodenum, adenocarcinoma"
                     /clone_lib="NIH_MGC_88"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             325. .1350
                     /codon_start=1
                     /product="MO25 protein"
                     /protein_id="AAH20570.1"
                     /db_xref="GI:18088261"
                     /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                     KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                     NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                     LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                     HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                     VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                     RPAQQEA"
BASE COUNT       1171 a     709 c     808 g    1073 t
ORIGIN
```

  Query Match   57.5%; Score 582.6; DB 9; Length 3761;
  Best Local Similarity   74.7%; Pred. No. 4.3e-132;
  Matches 748; Conservative   0; Mismatches 244; Indels   9; Gaps   1;

```
Qy       18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
            ||||| | || || |||||| ||||||||| |||||||| |  ||||| || |   ||||
Db      336 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 395

Qy       78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
            || ||||||||||||||||      ||||| |||| ||||||| |||||||||| ||
Db      396 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 455

Qy      129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
            |||   |||   || ||||||||||||||||||| ||| ||||| || ||||| || | ||
Db      456 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 515

Qy      189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
            |||||||| |||| || || ·||||||||| | ||||| || || || ||| |
Db      516 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 575

Qy      249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
```

```
        ||||||   |  ||||||  ||  ||||||||||||  |||||||||  |||  |  ||  ||  ||  |||||
Db   576 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 635

Qy   309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
         ||  |  |||||||||  ||  ||  ||  |  |  ||||||||  ||  ||  ||  |  |||
Db   636 TATTCTCGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 695

Qy   369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
         ||||   ||||  |||  |   |  |||||  ||||||  |  |||  |  ||  ||  ||||||
Db   696 GAATATTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 755

Qy   429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
         ||  |||  |  |||||||||  ||  ||||||||||||||||||  |||||||||  |  |  ||  |
Db   756 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 815

Qy   489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
         ||  ||   ||||||||  |  ||  ||  ||  ||||||||||||  ||  |||||||||||||
Db   816 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 875

Qy   549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
         ||||| ||  ||||||||||||||||||  ||  ||||||||||  ||  |      |||||  ||  ||  ||
Db   876 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 935

Qy   609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
         |||  ||||  ||  |     ||||    ||  |||||||||  ||  |||||  ||  ||  |||||||||
Db   936 ACAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 995

Qy   669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
         ||  ||  |||||||||  |  |||||  ||  ||  ||  ||  |  ||  ||  |  ||||||||  |
Db   996 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 1055

Qy   729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
         ||  ||||||||||  ||  ||||||  ||  ||  ||||||||||  |||  |  |||||||||||  ||  ||
Db  1056 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1115

Qy   789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
         ||  ||||||||  |||||||||||||||||  |||||||||  ||||||||||||||||||||  ||||
Db  1116 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 1175

Qy   849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
         |||| ||||  ||  ||||||  ||  |  ||  ||||||  |  ||  ||  |||  |||||||||||  ||
Db  1176 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1235

Qy   909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
         ||| ||  ||||  ||  ||  ||  ||  ||||||||  |||||||||||         ||||||||||
Db  1236 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1295

Qy   969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
         ||| ||  ||||||||||  |  ||  ||||||||  |  |  ||  |
Db  1296 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1336


RESULT 14
AF151824
LOCUS       AF151824                 1680 bp    mRNA    linear    PRI 18-MAY-2000
DEFINITION  Homo sapiens CGI-66 protein mRNA, complete cds.
ACCESSION   AF151824
```

```
VERSION        AF151824.1  GI:4929600
KEYWORDS       .
SOURCE         Homo sapiens (human)
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 1680)
  AUTHORS      Lai,C.H., Chou,C.Y., Ch'ang,L.Y., Liu,C.S. and Lin,W.
  TITLE        Identification of novel human genes evolutionarily conserved in
               Caenorhabditis elegans by comparative proteomics
  JOURNAL      Genome Res. 10 (5), 703-713 (2000)
  MEDLINE      20272150
   PUBMED      10810093
REFERENCE      2  (bases 1 to 1680)
  AUTHORS      Lin,W.-C.
  TITLE        Direct Submission
  JOURNAL      Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia
               Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan
FEATURES             Location/Qualifiers
     source          1. .1680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             1. .1026
                     /codon_start=1
                     /product="CGI-66 protein"
                     /protein_id="AAD34061.1"
                     /db_xref="GI:4929601"
                     /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                     KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                     NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                     LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                     HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                     VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                     RPAQQEA"
BASE COUNT       540 a    324 c    358 g     458 t
ORIGIN

   Query Match                57.4%;  Score 581.6;  DB 9;  Length 1680;
   Best Local Similarity      74.7%;  Pred. No. 7.5e-132;
   Matches  747; Conservative    0; Mismatches  244; Indels    9; Gaps    1;

Qy      19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
           ||| | || || ||||||| |||||||||| ||||||||| |  ||||| || |   ||||
Db      13 TTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGCT 72

Qy      79 ATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTCT 129
           || |||||||||||||||         |||||  |||| |||||| |||||||||||| ||
Db      73 GTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTCC 132

Qy     130 AAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACA 189
           |||    ||| || |||||||||||||||||| ||| ||||| || ||||| || |  |||
Db     133 AAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAGAGCCTCAGACA 192

Qy     190 GAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATA 249
           |||||||| ||||| || || ||||||||||| | ||||| || || || || ||| ||
Db     193 GAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGTA 252
```

```
Qy   250 GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAAC 309
         |||||  |  |||||| || |||||||||||| ||||||||| ||| | || || || |||||
Db   253 GCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAAT 312

Qy   310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCT 369
         ||  |  |||||||||  || || ||   |  |  |||||||||| || || |   | ||| 
Db   313 ATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACAG 372

Qy   370 CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG 429
          ||||    ||||  ||| |||  |  | |||||| |||||||  | ||| | || || ||||||
Db   373 AATATTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGGA 432

Qy   430 ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
         || ||| | |||||||| ||   |||||||||||||||| ||||||| ||  |  | || |
Db   433 ATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGAA 492

Qy   490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
         || ||   |||| ||||| |  || || || ||  ||||||||||| || ||||||||||||
Db   493 CAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGCA 552

Qy   550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
         |||||  ||  ||||||||||||| || ||||||||| ||  |||||||||  |  | |||
Db   553 TTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGAA 612

Qy   610 CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
         ||  |||| || |     ||||   || |||||||| || ||||| || ||  |||||||||
Db   613 CAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTG 672

Qy   670 ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
         || ||  |||||||| | | ||||| || || || || || ||  | || || |||||||| |
Db   673 ACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACA 732

Qy   730 ATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG 789
         || ||||||||| || ||||| || || |||||||| ||| | |||||||||||||| || ||
Db   733 ATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGA 792

Qy   790 GATAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGT 849
         || ||||||| ||||||||||||||| |||||| |||| ||||||||||||||| |||| |
Db   793 GACAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAAT 852

Qy   850 CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAG 909
         ||| ||||  || |||||| ||  || |  | || |||||  |  | || ||  |||||||||||| |||
Db   853 CCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAG 912

Qy   910 TTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC 969
         || || |||| || || || || || |||||||| |||||||||||    |||||||||| |
Db   913 TTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACC 972

Qy   970 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
         || ||  |||||||||||| | || |||||||| | || |
Db   973 TATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012
```

RESULT 15
AF113536
LOCUS   AF113536   3466 bp   mRNA   linear   PRI 04-DEC-1999

```
DEFINITION   Homo sapiens MO25 protein mRNA, complete cds.
ACCESSION    AF113536
VERSION      AF113536.1  GI:6523826
KEYWORDS     .
SOURCE       Homo sapiens (human)
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 3466)
  AUTHORS    Jin,W., Shi,J., Ren,S., Gu,J., Fu,S., Huang,Q., Dong,H., Yu,Y.,
             Fu,G., Wang,Y., Chen,Z. and Han,Z.
  TITLE      A novel gene expressed in the human hypothalamus
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 3466)
  AUTHORS    Jin,W., Shi,J., Ren,S., Gu,J., Fu,S., Huang,Q., Dong,H., Yu,Y.,
             Fu,G., Wang,Y., Chen,Z. and Han,Z.
  TITLE      Direct Submission
  JOURNAL    Submitted (16-DEC-1998) Chinese National Human Genome Center at
             Shanghai, 351 Guo Shoujing Rd., Zhangjiang Hi-Tech Park, Pudong,
             Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1..3466
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="hypothalamus"
     CDS             54..1079
                     /codon_start=1
                     /product="MO25 protein"
                     /protein_id="AAF14873.1"
                     /db_xref="GI:6523827"
                     /translation="MPFPFGKSHKSPADIVKNLKESMAVLEKQDISDKKAEKATEEVS
                     KNLVAMKEILYGTNEKEPQTEAVAQLAQELYNSGLLSTLVADLQLIDFEGKKDVAQIF
                     NNILRRQIGTRTPTVEYICTQQNILFMLLKGYESPEIALNCGIMLRECIRHEPLAKII
                     LWSEQFYDFFRYVEMSTFDIASDAFATFKDLLTRHKLLSAEFLEQHYDRFFSEYEKLL
                     HSENYVTKRQSLKLLGELLLDRHNFTIMTKYISKPENLKLMMNLLRDKSRNIQFEAFH
                     VFKVFVANPNKTQPILDILLKNQAKLIEFLSKFQNDRTEDEQFNDEKTYLVKQIRDLK
                     RPAQQEA"
BASE COUNT     1101 a    606 c    689 g   1070 t
ORIGIN
```

```
     Query Match          57.4%;   Score 581.6;  DB 9;   Length 3466;
     Best Local Similarity  74.7%;   Pred. No. 7.6e-132;
     Matches  747;  Conservative    0;  Mismatches  244;  Indels   9;  Gaps     1;

Qy      19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
           ||| | || || |||||||  |||||||| |||||||| |  ||||| || |    ||||
Db      66 TTTGGAAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGCT 125

Qy      79 ATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTCT 129
           || ||||||||||||||        |||||  |||| ||||||| |||||||||| ||
Db     126 GTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTCC 185

Qy     130 AAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACA 189
           |||    |||   ||  || |||||||||||||||| ||| ||||| || ||||| || |   |||
Db     186 AAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGACA 245
```

```
Qy    190 GAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATA 249
          |||||||| |||||| || || ||||||||||| |  ||||| || ||    || ||| ||
Db    246 GAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGTA 305

Qy    250 GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAAC 309
          |||||  | ||||| || |||||||||| |||||||| ||| || || || |||||||
Db    306 GCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAAT 365

Qy    310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCT 369
          ||  |  |||||||| || || || |  | | ||||||||| || || ||  |  | || |
Db    366 ATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACAG 425

Qy    370 CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG 429
          ||||    |||| |||  |  | |  |||||| || || |  |  || ||    || |||||
Db    426 AATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGGA 485

Qy    430 ATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
          ||  |||  | ||||||||| ||  ||||||||||||| |||||||||| |  |  || |
Db    486 ATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTGTGGTCGGAA 545

Qy    490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
          || || ||||||||| |  || || || ||  ||||||||||||| || ||||||||||||
Db    546 CAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGCA 605

Qy    550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
          |||||  || ||||||||||||| || |||||||| || || |    ||||| || || |||
Db    606 TTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGAA 665

Qy    610 CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
          ||  |||| || |    ||||    || |||||||| || ||||| || || |||||||||
Db    666 CAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTG 725

Qy    670 ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
          || || |||||||  |||| | |||| || ||  |||  | |  |  || | | ||||||||  |
Db    726 ACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACA 785

Qy    730 ATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG 789
          || |||||||| || |||||| || || || |||||||| ||| | |||||||||||| || ||
Db    786 ATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGA 845

Qy    790 GATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGT 849
          || ||||||| ||||||||||||||||||||||| ||||||||||||||||||||  ||| |
Db    846 GACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAAT 905

Qy    850 CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAACTCATTGAG 909
          ||| |||| || ||||||| || |  || |||||  |  || || || ||| |||||||| |||
Db    906 CCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAG 965

Qy    910 TTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC 969
          || || |||| || || || || || ||||||||| |||||||||||||  |||||||||| |
Db    966 TTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACC 1025

Qy    970 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
          || || |||||||||||| ||| || ||||||||| |  || |
Db   1026 TATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1065
```

Search completed: January 6, 2004, 02:34:57
Job time : 3971 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 00:31:27 ; Search time 390 Seconds
                                               (without alignments)
                                               7018.539 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt.........tgaagaaaacggccccttga 1014

Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_19Jun03:*
                  1:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                  2:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                  3:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                  4:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                  5:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                  6:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                  7:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                  8:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                  9:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                  10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                  11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                  12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                  13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                  14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                  15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                  16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                  17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                  18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                  19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                  20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                  21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                  22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                  23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                  24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
                  25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1014	100.0	1014	22	AAF86462	Human Acute Neuron
	2	1014	100.0	1421	22	AAI58234	Human polynucleoti
	3	1010.8	99.7	1344	21	AAA27332	Human calcium bind
	4	992.8	97.9	2002	22	AAH15879	Human cDNA sequenc
	5	770.6	76.0	822	22	AAH05471	Human cDNA clone (
	6	684.6	67.5	831	20	AAX39817	Gastric cancer ass
c	7	684.4	67.5	1191	22	AAI60020	Human polynucleoti
	8	582.6	57.5	1026	22	AAC91772	Human ANIC-BP (acu
	9	582.6	57.5	3281	24	ABK13127	Human secretory po
	10	582.6	57.5	3849	23	ABV22987	Human prostate exp
	11	582.6	57.5	3849	23	ABV28822	Human prostate exp
	12	541.6	53.4	1053	22	AAF30688	Human acute neuron
	13	539.6	53.2	1162	23	AAS89557	DNA encoding novel
c	14	520.2	51.3	833	20	AAX39818	Gastric cancer ass
	15	496	48.9	2492	23	AAS88031	DNA encoding novel
	16	387.8	38.2	722	20	AAZ15133	Human gene express
	17	362.8	35.8	2231	23	ABL07151	Drosophila melanog
	18	362.8	35.8	4231	23	ABL07150	Drosophila melanog
	19	288.8	28.5	690	24	ABS77084	Frog embryonic gen
	20	246.4	24.3	435	24	ABL82285	Human ovarian canc
	21	244.8	24.1	447	24	ABL82921	Human ovarian canc
	22	244.8	24.1	450	24	ABL81975	Human ovarian canc
	23	210.8	20.8	762	24	ABS76784	Frog embryonic gen
	24	210.4	20.7	1474	21	AAC32983	Arabidopsis thalia
	25	208.8	20.6	1497	21	AAC40181	Arabidopsis thalia
	26	200.2	19.7	918	21	AAC42766	Arabidopsis thalia
	27	200.2	19.7	1032	21	AAC48253	Arabidopsis thalia
c	28	195	19.2	387	24	ABN93983	Gene #481 used to
c	29	195	19.2	387	24	ABL66143	Lung cancer relate
	30	169.8	16.7	722	24	AAS61992	Porcine muscular s
	31	166.6	16.4	700	24	AAS61993	Porcine muscular s
	32	163.8	16.2	481	25	ABZ19574	Group III cDNA can
	33	163.4	16.1	300	20	AAZ14552	Human gene express
	34	161.2	15.9	1515	21	AAC50415	Arabidopsis thalia
	35	156	15.4	861	24	ABN98824	Arabidopsis thalia
c	36	153.4	15.1	737	23	AAS79449	DNA encoding novel
	37	147.2	14.5	464	21	AAC46721	Zea mays DNA fragm
	38	133.2	13.1	615	22	AAH07116	Human cDNA clone (
	39	107.6	10.6	1149	23	AAS88030	DNA encoding novel
	40	107.6	10.6	3279	23	AAS89559	DNA encoding novel
	41	65.6	6.5	432	24	ABN78107	Human ORF3054 cDNA
	42	65	6.4	487	22	AAI98879	Human excretory re
	43	65	6.4	487	22	AAI64066	Human bladder rela
	44	53.6	5.3	254	25	ABX31310	Human GDP-mannose
	45	43	4.2	447	21	AAC06449	Human secreted pro

RESULT 1
AAF86462
ID   AAF86462 standard; cDNA; 1014 BP.
XX
AC   AAF86462;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP, cDNA.
XX
KW   Human; cerebroprotective; neuroprotective; vulnerary; vaccine;
KW   gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW   stroke; acute head trauma; multiple sclerosis; spinal cord injury; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1014
FT   /*tag= a
FT   /product= "Human Acute Neuronal Induced Calcium Binding
FT   Protein, ANIC-BP"
XX
PN   WO200123552-A1.
XX
PD   05-APR-2001.
XX
PF   18-SEP-2000; 2000WO-EP09132.
XX
PR   24-SEP-1999; 99EP-0118848.
XX
PA   (MERE) MERCK PATENT GMBH.
XX
PI   Den Daas I, Duecker K;
XX
DR   WPI; 2001-308142/32.
DR   P-PSDB; AAB82090.
XX
PT   Novel human acute neuronal induced calcium binding polypeptide, and
PT   polynucleotides encoding them useful for diagnosing or treating stroke,
PT   acute head trauma, multiple sclerosis and spinal cord injury -
XX
PS   Claim 5; Page 40-41; 45pp; English.
XX
CC   The present sequence is the coding sequence for human Acute Neuronal
CC   Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
CC   protein are useful for treating stroke, acute head trauma, multiple
CC   sclerosis and spinal cord injury. ANIC-BP coding sequence and protein
CC   are also useful as vaccines for inducing an immunological response in a
CC   mammal.
XX
SQ   Sequence 1014 BP; 340 A; 205 C; 209 G; 260 T; 0 other;

      Query Match   100.0%; Score 1014; DB 22; Length 1014;
      Best Local Similarity 100.0%; Pred. No. 3.5e-272;

Matches 1014; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

Qy    61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240

Qy   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300

Qy   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360

Qy   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420

Qy   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480

Qy   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540

Qy   541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600

Qy   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660

Qy   661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720

Qy   721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780

Qy   781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
```

```
Qy    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900

Qy    901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960

Qy    961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
```

RESULT 2
AAI58234
```
ID   AAI58234 standard; cDNA; 1421 BP.
XX
AC   AAI58234;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 437.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39078.
XX
```

```
     Query Match                 100.0%;  Score 1014;  DB 22;  Length 1421;
     Best Local Similarity       100.0%;  Pred. No. 4e-272;
     Matches 1014;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;


Qy          1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276

Qy         61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336

Qy        121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396

Qy        181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456

Qy        241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516

Qy        301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576

Qy        361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636

Qy        421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
```

```
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     637  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC  696

Qy     481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  540
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     697  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  756

Qy     541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  600
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     757  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  816

Qy     601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  660
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     817  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  876

Qy     661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  720
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     877  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  936

Qy     721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  780
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     937  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  996

Qy     781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  840
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     997  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  1056

Qy     841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  900
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1057  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  1116

Qy     901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  960
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1117  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  1176

Qy     961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1014
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1177  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1230
```

RESULT 3
AAA27332
ID   AAA27332 standard; cDNA; 1344 BP.
XX
AC   AAA27332;
XX
DT   10-AUG-2000 (first entry)
XX
DE   Human calcium binding protein hCBP gene.
XX
KW   Human; calcium binding protein; cancer; inflammation; CBP;
KW   reproductive disorder; autoimmune disorder; developmental disorder;
KW   seizure disorder; immune disorder; infection; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers

FT   CDS   124..1134
FT   /*tag= a
FT   /product= "calcium binding protein"
XX
PN   WO200029580-A1.
XX
PD   25-MAY-2000.
XX
PF   12-NOV-1999;   99WO-US27027.
XX
PR   13-NOV-1998;   98US-0190965.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT, Guegler KJ, Corley NC, Gorgone GA;
XX
DR   WPI; 2000-387793/33.
DR   P-PSDB; AAY94247.
XX
PT   Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
PT   diagnosis, prevention and treatment of cancers, immune, developmental
PT   or reproductive disorders -
XX
PS   Claim 9; Fig 1; 72pp; English.
XX
CC   The present sequence is the human calcium binding protein hCBP gene. It
CC   was obtained by screening a coronary artery smooth muscle cDNA library,
CC   from which five overlapping nucleic acids were isolated and
CC   sequenced, and then expressed to give the protein. The protein and the
CC   gene encoding it are useful for the diagnosis and treatment of the
CC   following types of disorder: cancers (such as adenocarcinomas),
CC   reproductive disorders (such as infertility, ovulatory defects,
CC   endometriosis, disruptions of the oestrus and menstrual cycles,
CC   polycystic ovary syndrome and ovarian hyperstimulation), autoimmune
CC   disorders (such as benign prostatic hyperplasia and prostatitis),
CC   developmental disorders (such as Cushing's syndrome, muscular dystrophy
CC   and gonadal dysgenesis), hereditary neuropathies, seizure disorders,
CC   immune disorders (such as AIDS, allergies, anaemia, asthma,
CC   atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'
CC   disease, multiple sclerosis, psoriasis, rheumatoid arthritis,
CC   scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,
CC   bacterial, fungal, parasitic, protozoal and helminthic infections.
XX
SQ   Sequence 1344 BP; 450 A; 261 C; 280 G; 353 T; 0 other;

      Query Match   99.7%; Score 1010.8; DB 21; Length 1344;
      Best Local Similarity 99.8%; Pred. No. 3.1e-271;
      Matches 1012; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

```
Qy    121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy    181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          |||||  |||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db    304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy    241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423

Qy    301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy    421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

Qy    481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy    541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy    601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy    661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy    721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy    781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023

Qy    901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy    961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
```

RESULT 4
AAH15879
ID   AAH15879 standard; cDNA; 2002 BP.
XX
AC   AAH15879;
XX
DT   26-JUN-2001 (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:14407.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 14407; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and

```
   Query Match              97.9%;   Score 992.8;  DB 22;   Length 2002;
   Best Local Similarity    99.8%;   Pred. No. 3.8e-266;
   Matches  994;  Conservative     0;  Mismatches     2;  Indels    0;  Gaps     0;


Qy        19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60


Qy        79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120


Qy       139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180


Qy       199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240


Qy       259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300


Qy       319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360


Qy       379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420


Qy       439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480


Qy       499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540


Qy       559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600


Qy       619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
```

```
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601   GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy    679   CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661   CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy    739   AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721   AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780

Qy    799   CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 858
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781   CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 840

Qy    859   ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 918
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841   ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGC 900

Qy    919   AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901   AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960

Qy    979   AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            ||||||||||||||||||||||||||||||||||||
Db    961   AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
```

RESULT 5
AAH05471
ID   AAH05471 standard; cDNA; 822 BP.
XX
AC   AAH05471;
XX
DT   26-JUN-2001   (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:2306.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
```

PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 1; SEQ ID 2306; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 822 BP; 268 A; 164 C; 171 G; 216 T; 3 other;

      Query Match   76.0%; Score 770.6; DB 22; Length 822;
      Best Local Similarity   98.5%; Pred. No. 2.1e-204;
      Matches 798; Conservative   0; Mismatches   10; Indels   2; Gaps   2;

Qy   19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

Qy   79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

Qy   139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180

Qy   199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
             ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

Qy   259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318

```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

Qy    319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

Qy    379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

Qy    439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480

Qy    499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

Qy    559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600

Qy    619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy    679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy    739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
          ||||||||||||||||||||||||||| ||||||||||||| |||||||||| |||||
Db    721 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778

Qy    799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
          ||||||||||||||||||  | | |  | ||
Db    779 CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808
```

RESULT 6
AAX39817
ID   AAX39817 standard; DNA; 831 BP.
XX
AC   AAX39817;
XX
DT   02-JUL-1999   (first entry)
XX
DE   Gastric cancer associated gene.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.

PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
XX
PS   Claim 67; Page 558-559; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence 831 BP; 285 A; 165 C; 167 G; 209 T; 5 other;

   Query Match   67.5%;   Score 684.6; DB 20;   Length 831;
   Best Local Similarity   96.1%;   Pred. No. 1.9e-180;
   Matches   764; Conservative   0; Mismatches   23; Indels   8; Gaps   6;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   37 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   97 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 156

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db     157 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 216

Qy     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 276

Qy     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 336

Qy     301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 396

Qy     361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456

Qy     421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 516

Qy     481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     517 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 576

Qy     541 TCAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTTGGTAGC-AG 598
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||| ||
Db     577 TCAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTTGGTAGCAAG 636

Qy     599 ACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 658
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     637 ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 696

Qy     659 AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 716
           || |||||||||| ||||||||||||| ||||| |||||| |||||| |||| ||||| |||||||
Db     697 AGAAATTATGTTACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAACTGATTCTGGACCG 756

Qy     717 TCACAACTTTGCCATC-ATGACAAAGTATATCAGCAAGCC---GGAGAACCTGAAACTCA 772
           ||| ||||||||||| || |||| ||||| |||||| ||| | || |
Db     757 TCANAACTTTGCCATCAANGCAAAAGTTTATCAACAAGCCNGGGGAAACCGGAAACNCAA 816

Qy     773 TGATGAACCTCCTTC 787
           || |||||||||||
Db     817 GGAGGAACCTCCTTC 831
```

RESULT 7
AAI60020/c
ID   AAI60020 standard; cDNA; 1191 BP.
XX
AC   AAI60020;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4009.
XX

```
KW      Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW      peripheral nervous system; neuropathy; central nervous system; CNS;
KW      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW      chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW      leukaemia; ss.
XX
OS      Homo sapiens.
XX
PN      WO200153312-A1.
XX
PD      26-JUL-2001.
XX
PF      26-DEC-2000; 2000WO-US34263.
XX
PR      21-JAN-2000; 2000US-0488725.
PR      25-APR-2000; 2000US-0552317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI      Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI      Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR      WPI; 2001-442253/47.
DR      P-PSDB; AAM40864.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders
PT      such as central nervous system injuries -
XX
PS      Claim 1; SEQ ID NO 4009; 10078pp; English.
XX
CC      The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC      the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC      immunosuppressant and cytostatic activity. The polynucleotides are useful
CC      in gene therapy. A composition containing a polypeptide or polynucleotide
CC      of the invention may be used to treat diseases of the peripheral nervous
CC      system, such as peripheral nervous injuries, peripheral neuropathy and
CC      localised neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: Immune system suppression,
CC      Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukaemias and
CC      C.N.S disorders.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification.
XX
SQ      Sequence 1191 BP; 348 A; 261 C; 236 G; 346 T; 0 other;
```

Query Match   67.5%; Score 684.4; DB 22; Length 1191;
Best Local Similarity 99.9%; Pred. No. 2.6e-180;
Matches 685; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy      329 GCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 388
            ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1189 GCACTCGAAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 1130

Qy      389 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 448
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1129 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 1070

Qy      449 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTA 508
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1069 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTA 1010

Qy      509 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 568
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1009 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 950

Qy      569 TACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 628
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      949 TACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 890

Qy      629 TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAA 688
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      889 TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAA 830

Qy      689 AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCA 748
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      829 AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCA 770

Qy      749 GCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCC 808
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      769 GCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCC 710

Qy      809 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTA 868
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      709 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTA 650

Qy      869 TTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAA 928
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      649 TTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAA 590

Qy      929 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 988
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      589 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 530

Qy      989 GAGACTTGAAGAAAACGGCCCCTTGA 1014
            ||||||||||||||||||||||||||
Db      529 GAGACTTGAAGAAAACGGCCCCTTGA 504
```

RESULT 8
AAC91772
ID   AAC91772 standard; cDNA; 1026 BP.
XX

```
AC    AAC91772;
XX
DT    27-MAR-2001   (first entry)
XX
DE    Human ANIC-BP (acute neuronal induced calcium-binding protein) cDNA.
XX
KW    Human; acute neuronal induced calcium-binding protein; ANIC-BP;
KW    Mo25 homologue; HymA homologue; drug screening; stroke;
KW    acute head trauma; multiple sclerosis; spinal cord injury; vaccine;
KW    cerebroprotective; neuroprotective; ss.
XX
OS    Homo sapiens.
XX
PN    WO200078947-A1.
XX
PD    28-DEC-2000.
XX
PF    14-JUN-2000; 2000WO-EP05457.
XX
PR    22-JUN-1999;   99EP-0112024.
XX
PA    (MERE ) MERCK PATENT GMBH.
XX
PI    Den Daas I,  Fischer V,  Seyfried C,  Von Melchner L;
XX
DR    WPI; 2001-102721/11.
DR    P-PSDB; AAB48970.
XX
PT    Novel acute neuronal induced calcium binding protein, useful for
PT    treating acute head trauma, stroke, multiple sclerosis and spinal cord
PT    injury  -
XX
PS    Claim 5; Page 35-36; 50pp; English.
XX
CC    The invention relates to human acute neuronal induced calcium-binding
CC    protein (ANIC-BP) and to nucleic acid encoding it. The invention
CC    also relates to expression systems and recombinant host cells comprising
CC    ANIC-BP DNA, the recombinant production of ANIC-BP, antibodies specific
CC    for ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin
CC    Fc region, and methods of screening for modulators of ANIC-BP function.
CC    ANIC-BP has homology and structural similarity to HymA and Mo25 proteins.
CC    ANIC-BP proteins and nucleotides are useful for treating stroke and
CC    acute head trauma, multiple sclerosis and spinal cord injury. ANIC-BP
CC    proteins are useful in screening assays, for identifying membrane bound
CC    or soluble receptors, and also in vaccines. ANIC-BP nucleotides are
CC    useful as diagnostic reagents, as tools for tissue expression studies,
CC    for chromosome localisation studies, as genetic vaccines, and in
CC    the generation of transgenic animals. The present sequence represents
CC    cDNA encoding human ANIC-BP.
XX
SQ    Sequence 1026 BP; 359 A; 199 C; 203 G; 265 T; 0 other;

   Query Match              57.5%;  Score 582.6;  DB 22;  Length 1026;
   Best Local Similarity    74.7%;  Pred. No. 5.5e-152;
   Matches  748;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    1;

Qy        18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
```

```
              |||| |  || ||  ||||||  |||||||||| ||||||||  |  |||||| || |    ||||
Db      12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC  71

Qy      78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC  128
              || ||||||||||||||     |||||    |||| ||||||  |||||||||| ||
Db      72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC  131

Qy     129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC  188
              |||   |||   ||  ||||||||||||||||  |||  ||||| || |||||| || | ||
Db     132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAGAGCCTCAGAC  191

Qy     189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT  248
              |||||||| |||||  ||  || ||||||||||| |   |||||  || ||    || ||| |
Db     192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT  251

Qy     249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA  308
              ||||||   | |||||| ||  |||||||||||| |||||||||| |||  |  ||  || |||||
Db     252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA  311

Qy     309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC  368
              || | |||||||||  || ||  ||  |  | ||||||||| || || ||   |   ||| |
Db     312 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA  371

Qy     369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG  428
              ||||   |||| ||| ||   |  ||||| ||||||  |  |||  | ||    ||||||
Db     372 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG  431

Qy     429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA  488
              || ||| | |||||||| ||  |||||||||||||||||||| |||||||  | |  || |
Db     432 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA  491

Qy     489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC  548
              || ||       |||||||| | || |  ||   ||||||||||||| || |||||||||||
Db     492 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC  551

Qy     549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA  608
              ||||| ||  ||||||||||||||| ||||||| || |    |||||| || || ||
Db     552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA  611

Qy     609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT  668
              |||  ||||  || |    ||||     || |||||||||| || |||||| || || |||||||||
Db     612 ACAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT  671

Qy     669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC  728
              || || ||||||||| | ||||| || || || ||  | || ||  ||||||||| |
Db     672 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC  731

Qy     729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG  788
              || ||||||||| || ||||| || ||  | ||||||||| ||| | |||||||||||| || ||
Db     732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG  791

Qy     789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG  848
              || |||||||  |||||||||||||||| ||||||||| |||||||||||||||| || || ||||
Db     792 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA  851

Qy     849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAAATCAGCCCAAACTCATTGA  908
              |||| |||| || || ||||| ||  | || |||||| | || || || |||| |||||||||||| ||
```

```
Db      852 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 911

Qy      909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
            ||| || |||| || || || || ||||||||| ||||||||||   |||||||||
Db      912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 971

Qy      969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
            ||| || ||||||||||| | || ||||||| | | || |
Db      972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012
```

RESULT 9
ABK13127
ID   ABK13127 standard; cDNA; 3281 BP.
XX
AC   ABK13127;
XX
DT   09-APR-2002   (first entry)
XX
DE   Human secretory polynucleotide (sptm) cDNA (481257.3).
XX
KW   Signal peptide; transmembrane domain; human; sptm; ss; gene;
KW   481257.3; antiarteriosclerotic; antiatherosclerotic; antipsoriatic;
KW   antiinflammatory; cytostatic; anti-HIV; antiallergic; antidiabetic;
KW   nephrotropic; antigout; antithyroid; hepatotropic; neuroprotective;
KW   osteopathic; antirheumatic; antiarthritic; dermatological; cancer;
KW   immunosuppressive; antiulcer; ophthalmological; vulnerary; gout;
KW   anticonvulsant; cerebroprotective; nootropic; antiparkinsonian;
KW   virucide; antibacterial; cell proliferative disorder; arteriosclerosis;
KW   atherosclerosis; psoriasis; immune system disorder; inflammation;
KW   acquired immunodeficiency syndrome; AIDS; Addison's disease;
KW   adult respiratory distress syndrome; allergy; cirrhosis; osteoporosis;
KW   diabetes mellitus; Graves' disease; multiple sclerosis; osteoarthritis;
KW   rheumatoid arthritis; systemic lupus erythematosus; ulcerative colitis;
KW   haematopoietic cancer; neurological disorder; stroke; epilepsy;
KW   Huntington's disease; Parkinson's disease; meningitis; prion disease;
KW   kuru; Creutzfeldt-Jakob disease; cerebral palsy; myasthaenia gravis;
KW   diabetic neuropathy; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200111032-A1.
XX
PD   15-FEB-2001.
XX
PF   01-JUN-2000; 2000WO-US15246.
XX
PR   05-AUG-1999;   99US-147500P.
PR   05-AUG-1999;   99US-147501P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hodgson DM,   Lincoln SE,   Russo FD,   Spiro PA,   Banville SC;
PI   Bratcher SR,   Dufour GE,   Cohen HJ,   Rosen BH,   Chalup MS,   Hillman JL;
PI   Jones AL,   Yu JY,   Greenawalt LB,   Panzer SR,   Roseberry AM;
PI   Wright RJ,   Daniels SE;
XX
```

PT    Novel secretory polynucleotide (sptm) and polypeptides encoded by sptm,
PT    useful for diagnosing and treating disorders or diseases associated
PT    with cell signaling e.g., allergy, psoriasis, Grave's disease, epilepsy
PT    -
XX
PS    Claim 1; Page 192-193; 198pp; English.
XX
CC    This invention relates to novel cDNA molecules encoding isolated
CC    secretory polynucleotides (sptm) with similarity to signal peptide
CC    (SP) or transmembrane domain (TM) consensus sequences. The
CC    polynucleotide sequences of the invention are useful for producing
CC    sptm protein by recombinant techniques, the protein may be used to
CC    generate anti-sptm antibodies which may be used to analyse protein
CC    expression levels in different tissues. The sptm molecules are useful
CC    for diagnostic and therapeutic purposes e.g., to diagnose or treat a
CC    condition associated with cell signaling such as a cell proliferative
CC    disorders (e.g., arteriosclerosis, atherosclerosis, psoriasis, cancers),
CC    immune system disorders (e.g., inflammation, acquired immunodeficiency
CC    syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,
CC    allergies, cirrhosis, diabetes mellitus, gout, Graves' disease,
CC    multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid arthritis,
CC    systemic lupus erythematosus, ulcerative colitis and haematopoietic
CC    cancer), a neurological disorder (e.g., stroke, epilepsy, Huntington's
CC    disease, Parkinson's disease, meningitis, prion diseases including kuru,
CC    Creutzfeldt-Jakob disease, cerebral palsy, myasthenia gravis, diabetic
CC    neuropathy and Alzheimer's disease). Sptm sequences can be used to
CC    detect the presence of or quantifying the amount of sptm-related
CC    polynucleotide in a sample. The sptm polynucleotide is used to design
CC    probes useful in diagnostic assays carried out to detect or confirm
CC    conditions, disorders, or diseases associated with abnormal levels of
CC    sptm expression. Sptm, its fragments or oligonucleotides derived from
CC    sptm may be used as primers in amplification steps prior to
CC    hybridisation. The present sequence represents the human sptm (481257.3)
CC    cDNA sequence of the invention.
XX
SQ    Sequence 3281 BP; 1014 A; 601 C; 676 G; 990 T; 0 other;

    Query Match            57.5%;   Score 582.6;  DB 24;  Length 3281;
    Best Local Similarity  74.7%;   Pred. No. 9.3e-152;
    Matches  748;  Conservative    0;  Mismatches  244;  Indels    9;  Gaps    1;

Qy       18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
            |||| | || || |||||| ||||||||| |||||||| | ||||| || |   ||||
Db      101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 160

Qy       78 CATTTTGGAAAAGCAAGAC---------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
            || ||||||||||||||||         ||||| |||| ||||||| ||||||||| ||
Db      161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 220

Qy      129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
            |||    |||    || |||||||||||||||||||| ||| ||||| || ||||| || | ||
Db      221 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 280

Qy      189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
            ||||||||| |||||| || || |||||||||||| | ||||| || ||     || ||| |

```
Db     281 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 340

Qy     249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
           ||||||  | ||||||  ||  ||||||||||||  |||||||| |||  | ||  || ||  ||||||
Db     341 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 400

Qy     309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
           ||  | |||||||||  || || ||  | | ||||||||| || || ||  |  | ||| |
Db     401 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 460

Qy     369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
           ||||    |||| |||  |   ||||| ||||||| | ||| |  | ||  ||  |||||| 
Db     461 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 520

Qy     429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
           || ||| | |||||||| || |  ||||||||||||||||||||| ||||||||  |  || |
Db     521 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTGTGGTCGGA 580

Qy     489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
           || ||    |||||||| |  || || || ||  ||||||||||| ||  ||||||||||||
Db     581 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 640

Qy     549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
           |||||  || |||||||||||||| ||  || |||||||||  ||  |    ||||| || || ||
Db     641 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 700

Qy     609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
           |||  ||||  ||  |   |||| |||||| | |||||||||| || |||||  ||  ||||||||
Db     701 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 760

Qy     669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
           || || |||||||| | |||||  ||  ||  || ||  |  | || || | ||||||||| |
Db     761 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 820

Qy     729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
           || |||||||| ||  ||||| ||  || || ||||||||| ||| | |||||||||||| || ||
Db     821 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 880

Qy     789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
           || ||||||||| |||||||||||||||||| |||||||| |||||||||||||| ||||||||||
Db     881 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 940

Qy     849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
           |||| ||||| |||| || |||| || | || ||||| | || || ||| ||||||||||| ||
Db     941 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1000

Qy     909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
           |||| || |||| || || || || || |||||||||| |||||||||||||||      ||||||||||
Db    1001 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1060

Qy     969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
           ||| || ||||||||||||| | || ||||||| | | ||| |
Db    1061 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1101
```

RESULT 10
ABV22987

```
ID    ABV22987 standard; cDNA; 3849 BP.
XX
AC    ABV22987;
XX
DT    13-SEP-2002  (first entry)
XX
DE    Human prostate expression marker cDNA 22978.
XX
KW    Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW    pharmacogenomic marker; gene; ss.
XX
OS    Homo sapiens.
XX
PN    WO200160860-A2.
XX
PD    23-AUG-2001.
XX
PF    20-FEB-2001; 2001WO-US05171.
XX
PR    17-FEB-2000; 2000US-183319P.
PR    16-MAR-2000; 2000US-189862P.
PR    25-MAY-2000; 2000US-207454P.
PR    09-JUN-2000; 2000US-211314P.
PR    18-JUL-2000; 2000US-219007P.
PR    13-DEC-2000; 2000US-255281P.
XX
PA    (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI    Schlegel R,  Endege WO,  Monahan JE;
XX
DR    WPI; 2001-662795/76.
XX
PT    Novel isolated nucleic acid molecule associated with cancerous state of
PT    prostate cells and correlating with presence of prostate cancer, useful
PT    for detecting presence of prostate cancer, stage of prostate cancer  -
XX
PS    Claim 1; Page 4088; 11750pp; English.
XX
CC    The invention relates to an isolated nucleic acid molecule (I) comprising
CC    a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC    specification or its complement. (I) is useful for:
CC    (a) assessing whether a patient is afflicted with prostate cancer;
CC    (b) monitoring the progression of prostate cancer in a patient;
CC    (c) assessing the efficacy of a test compound to inhibit prostate
CC    cancer in a patient;
CC    (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC    in a patient;
CC    (e) selecting a composition for inhibiting prostate cancer in a patient;
CC    (f) assessing the prostate cell carcinogenic potential of a compound;
CC    (g) determining whether prostate cancer has metastasized in a patient;
CC    (h) assessing the aggressiveness or indolence of prostate cancer in a
CC    patient;
CC    (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ    Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;
      Query Match            57.5%;  Score 582.6;  DB 23;  Length 3849;
```

Best Local Similarity   74.7%; Pred. No. 1e-151;
     Matches 748; Conservative   0; Mismatches 244; Indels   9; Gaps   1;

```
Qy      18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
           |||| | || || |||||||  |||||||||| ||||||||| |  ||||| || |  -||||
Db     437 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 496

Qy      78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
           || |||||||||||||||        |||| |||| ||||||| |||||||||| ||
Db     497 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 556

Qy     129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
           |||   |||   || ||||||||||||||||| ||| |||||| || ||||| || | ||
Db     557 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAGAGCCTCAGAC 616

Qy     189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
           ||||||||| |||||| || || |||||||||||| |  |||||| || || |  || ||| |
Db     617 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 676

Qy     249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
           ||||||  | ||||| || |||||||||||| |||||||| ||| | | || || || |||||
Db     677 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 736

Qy     309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
           ||  |  |||||||||| || || ||  | | ||||||||| || || || |  |  ||| |
Db     737 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 796

Qy     369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
           |||| |||| ||| |  | ||||| |||||| | ||| | || || || | |||||||
Db     797 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 856

Qy     429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
           || ||| | ||||||||| || |||||||||||||||||||| ||||||| | | | || |
Db     857 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTGTGGTCGGA 916

Qy     489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
           || ||      |||||||| | || || || || || |||||||||||| || |||||||||||
Db     917 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 976

Qy     549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
           |||| || ||||||||||||||||||  || |||||||||| || |  ||||| || || ||
Db     977 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 1036

Qy     609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
           ||| |||| || |  |||| |  || |||||||| || |||| | || | || |||||||||
Db    1037 ACAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 1096

Qy     669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
           || || ||||||||| |  || |||||| || || || || |  | ||| | |||||||| |
Db    1097 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 1156

Qy     729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
           || |||||||| || ||||| || || ||||||||| |||| | ||||||||||| || ||
Db    1157 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1216

Qy     789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
           || |||||||| ||||||||||||||||||| |||||||| ||||||||||||||||| ||||
```

```
Db      1217 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 1276

Qy       849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
              ||||  |||||  || |||||| ||   | || ||||||  | || ||  ||| |||||||||||| ||
Db      1277 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1336

Qy       909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
               ||| || ||||   || || || || |||||||| |||||||||||       ||||||||||
Db      1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1396

Qy       969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
              ||| ||  ||||||||||| | || |||||||| | | || |
Db      1397 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1437
```

RESULT 11
ABV28822
ID   ABV28822 standard; cDNA; 3849 BP.
XX
AC   ABV28822;
XX
DT   16-SEP-2002   (first entry)
XX
DE   Human prostate expression marker cDNA 28813.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Endege WO, Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS   Claim 1; Page 6066-6067; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;

     Query Match   57.5%; Score 582.6; DB 23; Length 3849;
     Best Local Similarity   74.7%; Pred. No. 1e-151;
     Matches 748; Conservative   0; Mismatches 244; Indels   9; Gaps   1;

Qy   18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
             |||| | || || |||||| |||||||||| |||||||| | ||||| || | ||||
Db   437 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 496

Qy   78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
             || |||||||||||||||| ||||| |||| |||||| ||||||||||| ||
Db   497 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 556

Qy   129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
             ||| ||| || ||||||||||||||||||| ||| ||||| || ||||| || | ||
Db   557 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 616

Qy   189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
             |||||||||| |||||| || || |||||||||||| | |||||| || || || ||| |
Db   617 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 676

Qy   249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
             |||||| | ||||| || |||||||||| ||||||||| ||| | || || || |||||
Db   677 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 736

Qy   309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
             || | ||||||||| || || || | | |||||||||| || || || | | || |
Db   737 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 796

Qy   369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
             |||| |||| ||| | | ||||| |||||| | ||| | || || || || ||||||
Db   797 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 856

Qy   429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
             || ||| | ||||||||| || |||||||||||||||||||| | | || | |
Db   857 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 916

Qy   489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
             || || ||||||||| | || || || ||||||||||||||| || ||||||||||||
Db   917 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 976

Qy   549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608

```
                  |||||  ||  |||||||||||||| || |||||||||| || |     ||||| || || ||
Db    977 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 1036

Qy    609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
                  |||  |||| || |    ||||    || ||||||||| || |||||| || || |||||||||
Db   1037 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 1096

Qy    669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
                  || || |||||||||  |  ||||| || || || ||   | ||| ||  | ||||||||  |
Db   1097 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 1156

Qy    729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
                  ||  |||||||| || |||||  || ||  |||||||||  |||  |  ||||||||||| || ||
Db   1157 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1216

Qy    789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
                  || ||||||||  |||||||||||||||||  ||||||||||  ||||||||||||||||||| . ||||
Db   1217 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 1276

Qy    849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
                  ||||  |||| || ||||| ||   | || |||||  | || || ||| |||||||||||| ||
Db   1277 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1336

Qy    909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
                  ||| || ||||  || || || || || ||||||||| |||||||||||      |||||||||||
Db   1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1396

Qy    969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
                  ||| ||  |||||||||||| | || ||||||| |  | || |
Db   1397 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1437


RESULT 12
AAF30688
ID    AAF30688 standard; cDNA; 1053 BP.
XX
AC    AAF30688;
XX
DT    11-JUN-2001  (first entry)
XX
DE    Human acute neuronal induced calcium binding protein ANIC-BP-1B cDNA.
XX
KW    Acute neuronal induced calcium binding protein; ANIC-BP-1B;
KW    spice variant; human; stroke; head trauma; Parkinson's disease;
KW    Alzheimer's disease; multiple sclerosis; spinal cord injury;
KW    cerebroprotective; antiparkinsonian; nootropic; neuroprotective;
KW    therapy; diagnosis; vaccine; ss.
XX
OS    Homo sapiens.
XX
FH    Key             Location/Qualifiers
FT    CDS             1..1053
FT                    /*tag=  a
FT                    /product= "Human ANIC-BP-1B"
XX
PN    WO200125423-A1.
XX
```

PT   Novel human acute neuronal induced calcium-binding protein like protein
PT   splice variant, useful for treating stroke, acute head trauma,
PT   Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal
PT   cord injury -
XX
PS   Claim 4; Page 43-44; 49pp; English.
XX
CC   The present sequence is that of cDNA encoding a novel human acute
CC   neuronal induced calcium binding protein-like protein splice
CC   variant, ANIC-NP-1B (see AAB20387). The protein shows homology to
CC   other members of the calcium binding protein family, including
CC   ANIC-BP, a protein discovered by mRNA differential display that is
CC   upregulated in a rat model of head trauma. ANIC-BP and ANIC-BP-1B
CC   differ in their C-terminal portions. The variant protein could
CC   serve as a novel drug target. The invention provides ANIC-BP-1B
CC   polynucleotides and polypeptides, expression vectors, host cells
CC   and antibodies, as well as methods for producing the protein and
CC   for treating or preventing disorders associated with expression of
CC   the protein by inhibiting or activating the action of ANIC-BP-1B.
CC   Diseases that may be treated include stroke and acute head trauma,
CC   Parkinson's disease, Alzheimer's disease, multiple sclerosis and
CC   spinal cord injury. The polynucleotides and polypeptides can also
CC   be used in diagnostic assays and in vaccines, and to identify
CC   agonists and antagonists useful for treating conditions associated
CC   with ANIC-BP-1B imbalance.
XX
SQ   Sequence 1053 BP; 357 A; 211 C; 214 G; 271 T; 0 other;

   Query Match   53.4%;   Score 541.6; DB 22;   Length 1053;
   Best Local Similarity   74.1%;   Pred. No. 1.5e-140;
   Matches   716; Conservative   0; Mismatches 239; Indels   11; Gaps   2;

Qy   18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
               ||||| | || || ||||||| ||||||||||| |||||||| | ||||| || | ||||
Db   12 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 71

Qy   78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
                 || ||||||||||||||| ||||| |||| ||||||| |||||||||| ||
Db   72 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 131

Qy   129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
                ||| ||| || ||||||||||||||||| ||| ||||| || ||||| || | | ||
Db   132 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC 191

```
Qy   189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
         ||||||||| ||||| || || |||||||||||| | ||||| || ||    || ||| |
Db   192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCTGGT 251

Qy   249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
         |||||| | |||||| || |||||||||||| |||||||| ||| | || || || |||||
Db   252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 311

Qy   309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
         || | ||||||||| || || || | | |||||||||| || || | | | | ||| 
Db   312 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 371

Qy   369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
          ||||  |||| ||| | |  ||||| |||||| | ||| | || ||  |||||||
Db   372 GAATATTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 431

Qy   429 GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
         || ||| | |||||||| || |||||||||||||||||| |||||||| || | || |
Db   432 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 491

Qy   489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
         || ||       |||||||| | || || || ||||||||||||| || |||||||||||
Db   492 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 551

Qy   549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
         |||| || |||||||||||||| || |||||||||| || |    ||||| || || ||
Db   552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTTGGA 611

Qy   609 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
         ||| |||| || |  |||| || |||||||| || ||||| || || |||||||
Db   612 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 671

Qy   669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
         || || ||||||||  || | ||||| || || || || ||  | || ||  |||||||| |
Db   672 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCAC 731

Qy   729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
         || |||||||| || || ||||| || || || ||||||||| || | |||||||||| || ||
Db   732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 791

Qy   789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
         || |||||||| ||||||||||||||||| ||||||| || ||||||||||||||| ||||
Db   792 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 851

Qy   849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
         |||| |||| || |||||| ||  | || || ||||| | || || ||| |||||||||||| ||
Db   852 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 911

Qy   909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGAT--GATGAGCAGTTCGCTGACGAGAAG 966
         ||| || |||| || || || || |||||||||  |||||||||| | ||| |
Db   912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGATTGTATGAGCAGTTCCGTACCGACGAC 971

Qy   967 AACTAC 972
         | | |
Db   972 GAATTC 977
```

RESULT 13
AAS89557
ID   AAS89557 standard; cDNA; 1162 BP.
XX
AC   AAS89557;
XX
DT   13-FEB-2002   (first entry)
XX
DE   DNA encoding novel human diagnostic protein #25361.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C,   Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG25370.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 25361; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed

XX
SQ   Sequence 1162 BP; 383 A; 241 C; 258 G; 280 T; 0 other;

    Query Match   53.2%;   Score 539.6;   DB 23;   Length 1162;
    Best Local Similarity 73.9%;   Pred. No. 5.6e-140;
    Matches 743; Conservative   0; Mismatches 249; Indels 14; Gaps   4;

```
Qy        18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
              |||| | || || |||||| |||||||||| |||||||| | ||||| || |   ||||
Db       143 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 202

Qy        78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
              || ||||||||||||||||      ||||| |||| ||||||| |||||||||| ||
Db       203 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 262

Qy       129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
              |||    |||    ||  |||||||||||||||| ||| ||||| || ||||| || | ||
Db       263 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGATCCTCAGAC 322

Qy       189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
              |||||||| ||||| || || |||||||||||| |  ||||| || ||  |  || ||| |
Db       323 AGAAGCAGGAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTATCACCCTGGT 382

Qy       249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
              ||||||  | |||||| || ||||||||||||||||| ||| | || || || ||||||
Db       383 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 442

Qy       309 CATCTTGAGAAGACAGATAGGCA-CTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATC 367
              ||   | ||||||||| || || |    | | || ||||||||| || || ||   | ||
Db       443 TATTCTCAGAAGACAAATTGGTACCGAGAACTCCTACTGTTGAATACATCTGCACCCAAA 502

Qy       368 CTCA--TATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATT--GCCTTACGT 423
              |  |       |  |||| ||| |  | |||||| |||||| | || | |      || | |
Db       503 CAGAATATTTTTGTTCATGTTATTGAAAGGGTATGAATCTCCCAGAAATAGCTCTAAATT 562

Qy       424 TGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTT 483
              ||||| || ||| | |||||||| || |||||||||| |||||||||||||| || | |
Db       563 TGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGGCAAAATCATTTTGTGG 622

Qy       484 TCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCA 543
              ||   | || ||      |||||||| |  || ||  |||||||||||||| || ||||||
Db       623 TCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCA 682

Qy       544 GATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTC 603
              ||||| |||||| || ||||||||||||| |||||||||||| || || |   ||||| ||
Db       683 GATGCATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTT 742

Qy       604 TTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAAT 663
              || |||||| |||| || |     |||| || |||||||||||| || ||||| || | ||
Db       743 TTGGAACAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAAT 802

Qy       664 TATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAAC 723
              |||||| || || |||||||| | ||||| || || || ||  | ||| ||  | |||||
Db       803 TATGTGACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAAC 862
```

```
Qy     724 TTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTC 783
           ||  |  ||  ||||||||  ||  ||||||  ||  ||  ||||||||  |||  | |||||||||||
Db     863 TTCACAATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTG 922

Qy     784 CTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTG 843
           ||  ||  ||  ||||||||  |||||||||||||||||||  |||||||||  |||||||||||||||||||
Db     923 CTGCGAGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTA 982

Qy     844 GCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTC 903
           ||||  ||||  ||||  ||  |||||  ||  |  ||  |||||  |  ||  ||  |||  |||||||||
Db     983 GCCAATCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTC 1042

Qy     904 ATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAG 963
           ||  |||||  ||  ||||  ||  ||  ||  ||  ||  |||||||||  |||||||||||        ||||||
Db    1043 ATAGAGTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAG 1102

Qy     964 AAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
           ||||  |||  ||  ||||||||||||  |  ||  ||||||||  |  |  ||  |
Db    1103 AAGACCTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1148


RESULT 14
AAX39818/c
ID     AAX39818 standard; DNA; 833 BP.
XX
AC     AAX39818;
XX
DT     02-JUL-1999   (first entry)
XX
DE     Gastric cancer associated gene.
XX
KW     Cancer associated antigen; diagnosis; research; treatment; human;
KW     breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW     prostate cancer; ss.
XX
OS     Homo sapiens.
XX
PN     WO9904265-A2.
XX
PD     28-JAN-1999.
XX
PF     15-JUL-1998;   98WO-US14679.
XX
PR     22-JUN-1998;   98US-0102322.
PR     17-JUL-1997;   97US-0896164.
PR     10-OCT-1997;   97US-0061599.
PR     10-OCT-1997;   97US-0061765.
PR     10-OCT-1997;   97US-0948705.
PR     11-OCT-1997;   97GB-0021697.
XX
PA     (LUDW-) LUDWIG INST CANCER RES.
XX
PI     Chen Y,  Gout I,  Gure A,  O'Hare M,  Obata Y,  Old LJ;
PI     Pfreundschuh M,  Sahin U,  Scanlan MJ,  Stockert E;
PI     Tureci O;
XX
DR     WPI; 1999-132448/11.
```

CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence 833 BP; 253 A; 171 C; 172 G; 227 T; 10 other;

  Query Match   51.3%;   Score 520.2;   DB 20;   Length 833;
  Best Local Similarity   98.1%;   Pred. No. 1.2e-134;
  Matches   566; Conservative   0; Mismatches   7; Indels   4; Gaps   4;

Qy   442 GAATGTATTCGACATGAACCACTTG-CCAAAATCATCCTC-TTTTCTAATCAATTCAGAG 499
            |||| |||||||||| ||| ||| ||| ||||| |||||||| |||||||||||||||||||
Db   732 GAATNTATTCGACTTGACCCANTTGCCCAAANTCATCCTCTTTTTCTAATCAATTCAGAG 673

Qy   500 ATTTCTTTAAGT-ACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   672 ATTTCTTTAAGTAACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 613

Qy   559 TTCAAGGATTTACTAACCAGA-CATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA 617
            ||||||||||||||||||| || | |||||||||||||||||||||||||||||||||||
Db   612 TTCAAGGATTTACTAACCNGACCTTAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA 553

Qy   618 CGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG 677
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 CGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG 493

Qy   678 ACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGAC 737
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   492 ACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGAC 433

Qy   738 AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG 797
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   432 AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG 373

Qy   798 TCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAA 857
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 TCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAA 313

Qy   858 AACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAG 917

```
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 AACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAG 253

Qy     918 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGAT 977
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGAT 193

Qy     978 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              |||||||||||||||||||||||||||||||||||||
Db     192 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 156
```

RESULT 15
AAS88031
ID   AAS88031 standard; cDNA; 2492 BP.
XX
AC   AAS88031;
XX
DT   13-FEB-2002 (first entry)
XX
DE   DNA encoding novel human diagnostic protein #23835.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG23844.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 23835; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
```

CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2492 BP; 751 A; 477 C; 546 G; 718 T; 0 other;

     Query Match   48.9%; Score 496; DB 23; Length 2492;
     Best Local Similarity   73.0%; Pred. No. 1.1e-127;
     Matches 737; Conservative   0; Mismatches 255; Indels 18; Gaps   7;

Qy   18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
             |||| | || || |||||| |||||||||| ||||||||| | ||||| || | ||||
Db   143 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 202

Qy   78 CATTTTGGAAAAGCAAGAC--------AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
              || ||||||||||||||| ||||| |||| |||||| |||||||||| ||
Db   203 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 262

Qy   129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC 188
             ||| ||| || ||||||||||||||||| ||| ||||| || ||||| || ||
Db   263 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGATCCTCAGAC 322

Qy   189 AGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
             |||||||| ||||| || || ||||||||||| | |||||| || || | || ||| |
Db   323 AGAAGCAGGAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTATCACCCTGGT 382

Qy   249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAA 308
             |||||| | |||||| || |||||||||||||||| |||||||| ||| | || || || |||||
Db   383 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 442

Qy   309 CATCTTGAGAAGACAGATAGGCA-CTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATC 367
             || | ||||||||| || || | | | | ||||||||| || || || | | ||
Db   443 TATTCTCAGAAGACAAATTGGTACCGAGAACTCCTACTGTTGAATACATCTGCACCCAAA 502

Qy   368 CTCA--TATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATT--GCCTTACGT 423
             | | | |||| ||| | | |||||| |||||| | || | | ||| |
Db   503 CAGAATATTTTTGTTCATGTTATTGAAAGGGTATGAATCTCCCAGAAATAGCTCTAAATT 562

Qy   424 TGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTT 483
             ||||| || ||| | ||||||||| || |||||||||||||||||||||| |||||||| ||
Db   563 TGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGGCAAAATCATTTTGTGG 622

Qy   484 TCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCA 543
             || | || || |||||||| | || || || |||||||||||||| || ||||||
Db   623 TCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCA 682

```
Qy    544 GATGCC-TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTT 602
          |||||| |||||| || |||||||| |||||| || ||||||||| || |    ||||| ||
Db    683 GATGCCATTTGCCACATTCAAGGGTTTACTTACAAGACATAAATTGCTCAGTGCAGAATT 742

Qy    603 CTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAA 662
          || |||||  |||| || |    ||||     || ||||||||| || ||||| || || ||
Db    743 TTTGGAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAA 802

Qy    663 TTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAA 722
          |||||| || || |||||||| | |||||| || || || || | || || | |||||
Db    803 TTATGTGACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAA 862

Qy    723 CTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAG--AACCTGAAACTCATGATGAAC 780
          |||   | || |||||||| || ||||| || || | |  | || | |||||||||
Db    863 CTTCACAATTATGACAAAATACATCAGTAAACCTGTGGAACCTCAAATTTAATGATGAAC 922

Qy    781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGG-TGTT 839
          || || || || |||||||| ||||||||||||||| ||||||||| ||||||||||  |||
Db    923 CTGCTGCGAGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGCAGTT 982

Qy    840 TGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 899
          ||| |||| |||| |||| || ||||| || | || |||||| | | || ||| ||||
Db    983 TGTAGCCAATCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAA 1042

Qy    900 ACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGA 959
          |||||| ||||| || |||| || || || || ||||||||| ||||||||||    ||
Db   1043 ACTCATAGAGTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGA 1102

Qy    960 CGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
          |||||||| ||| || |||||||||||| | || ||||||| | | ||| |
Db   1103 CGAGAAGACCTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1152
```

Search completed: January 6, 2004, 01:28:41
Job time : 394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 02:35:04 ; Search time 1394 Seconds
                                           (without alignments)
                                           2517.743 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt..........tgaagaaacggccccttga 1014

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_NA:*
                1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
                14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                              SUMMARIES
                        %
Result   Query
  No.   Score Match Length DB   ID   Description

```
  --------------------------------------------------------------------------------
      1     1014   100.0    1421   13  US-10-117-722-111        Sequence 111, App
      2     1014   100.0    1421   15  US-10-037-270-111        Sequence 111, App
      3   1010.8    99.7    1344   15  US-10-025-730-2          Sequence 2, Appli
      4      398    39.3     475   11  US-09-918-995-5343       Sequence 5343, Ap
      5    288.8    28.5     690    9  US-09-910-943-318       Sequence 318, App
      6    246.4    24.3     435   10  US-09-867-701-5263       Sequence 5263, Ap
      7    244.8    24.1     447   10  US-09-867-701-5899       Sequence 5899, Ap
      8    244.8    24.1     450   10  US-09-867-701-4953       Sequence 4953, Ap
      9    210.8    20.8     762    9  US-09-910-943-35        Sequence 35, Appl
C    10      195    19.2     387   10  US-09-954-456-1453       Sequence 1453, Ap
C    11      195    19.2     387   10  US-09-880-107-481        Sequence 481, App
     12    169.8    16.7     722   13  US-10-257-826A-118       Sequence 118, App
     13    166.6    16.4     700   13  US-10-257-826A-119       Sequence 119, App
     14      156    15.4     861    9  US-09-770-445-592       Sequence 592, App
     15     74.2     7.3     262    9  US-09-923-876-1251      Sequence 1251, Ap
     16     74.2     7.3     262   12  US-09-923-876-1251       Sequence 1251, Ap
C    17     65.6     6.5     336   11  US-09-918-995-19069      Sequence 19069, A
     18       65     6.4     487   12  US-10-242-355-323        Sequence 323, App
     19       65     6.4     487   13  US-10-080-254-54         Sequence 54, Appl
     20     53.6     5.3     254   10  US-09-878-574-13369      Sequence 13369, A
C    21     50.6     5.0     486   11  US-09-770-961-777        Sequence 777, App
     22     41.6     4.1     242    9  US-09-923-876-2528      Sequence 2528, Ap
     23     41.6     4.1     242   12  US-09-923-876-2528       Sequence 2528, Ap
     24     40.8     4.0    1295   12  US-10-310-154-294        Sequence 294, App
     25     40.2     4.0  113306   12  US-10-292-798-1007       Sequence 1007, Ap
     26     39.8     3.9     431   11  US-09-918-995-5787       Sequence 5787, Ap
C    27     38.4     3.8    6301   13  US-10-311-455-26         Sequence 26, Appl
     28     38.2     3.8    1200   13  US-10-027-632-261235     Sequence 261235,
     29     38.2     3.8    1200   14  US-10-027-632-261235     Sequence 261235,
     30     37.8     3.7    1457   15  US-10-054-968-9          Sequence 9, Appli
     31     37.8     3.7    7178   13  US-09-873-367C-278       Sequence 278, App
     32     37.6     3.7    1267   14  US-10-001-843-45         Sequence 45, Appl
C    33     37.6     3.7 3673778   13   US-10-312-841-1          Sequence 1, Appli
     34     37.4     3.7    2232   15  US-10-087-464-45         Sequence 45, Appl
C    35     37.4     3.7    4012    9  US-09-876-889-335       Sequence 335, App
C    36     37.4     3.7    4103   13  US-10-117-722-390        Sequence 390, App
C    37     37.4     3.7    4103   15  US-10-037-270-390        Sequence 390, App
C    38     37.4     3.7    8577   13  US-10-311-455-1760       Sequence 1760, Ap
C    39     37.2     3.7     869   13  US-10-027-632-261978     Sequence 261978,
C    40     37.2     3.7     869   14  US-10-027-632-261978     Sequence 261978,
C    41       37     3.6    5413   13  US-10-311-455-538        Sequence 538, App
C    42     36.6     3.6    9367   13  US-10-311-455-944        Sequence 944, App
     43     36.4     3.6     461   14  US-10-079-623-143        Sequence 143, App
     44     36.4     3.6    2641   12  US-10-369-493-29299      Sequence 29299, A
C    45     36.4     3.6    6071   13  US-10-311-455-297        Sequence 297, App
```

ALIGNMENTS

RESULT 1
US-10-117-722-111
; Sequence 111, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom

```
    Query Match              100.0%;  Score 1014;  DB 13;  Length 1421;
    Best Local Similarity    100.0%;  Pred. No. 1.2e-281;
    Matches 1014;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      217 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276

Qy       61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
```

```
Qy      421  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      637  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC  696

Qy      481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      697  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  756

Qy      541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      757  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  816

Qy      601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      817  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  876

Qy      661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      877  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  936

Qy      721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      937  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  996

Qy      781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      997  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  1056

Qy      841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1057  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  1116

Qy      901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1117  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  1176

Qy      961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1014
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1177  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1230
```

RESULT 2
US-10-037-270-111
; Sequence 111, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
;   APPLICANT: Tang, Y. Tom
;   APPLICANT: Liu, Chenghua
;   APPLICANT: Asundi, Vinod
;   APPLICANT: Zhang, Jie
;   APPLICANT: Ren, Feiyan
;   APPLICANT: Chen, Rui-hong
;   APPLICANT: Zhao, Qing A.
;   APPLICANT: Wehrman, Tom
;   APPLICANT: Xue, Aidong J.
;   APPLICANT: Yang, Yonghong

```
   Query Match            100.0%;  Score 1014;  DB 15;  Length 1421;
   Best Local Similarity  100.0%;  Pred. No. 1.2e-281;
   Matches 1014;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      217 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276

Qy       61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 456

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 516

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       577  GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA  636

Qy       421  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       637  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC  696

Qy       481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       697  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT  756

Qy       541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       757  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC  816

Qy       601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       817  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG  876

Qy       661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       877  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC  936

Qy       721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       937  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC  996

Qy       781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       997  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT  1056

Qy       841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1057  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA  1116

Qy       901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1117  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC  1176

Qy       961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1014
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1177  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1230
```

RESULT 3
US-10-025-730-2
; Sequence 2, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
;   APPLICANT: Tang, Y. Tom
;   APPLICANT: Guegler, Karl J.
;   APPLICANT: Corley, Neil C.
;   APPLICANT: Gorgone, Gina A.
;   TITLE OF INVENTION: CALCIUM BINDING PROTEIN
;   FILE REFERENCE: PF-0635 US
;   CURRENT APPLICATION NUMBER: US/10/025,730
;   CURRENT FILING DATE: 2001-12-18
;   PRIOR APPLICATION NUMBER: US/09/190,965

```
    Query Match              99.7%;  Score 1010.8;  DB 15;  Length 1344;
    Best Local Similarity    99.8%;  Pred. No. 1e-280;
    Matches 1012;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy        1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy       61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            |||||  |||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db      304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 423

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

Qy      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       724  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy       661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       784  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy       721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       844  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy       781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       904  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy       841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       964  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 1023

Qy       901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1024  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy       961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1084  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137
```

RESULT 4
US-09-918-995-5343
; Sequence 5343, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5343
;   LENGTH: 475
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(475)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-5343

  Query Match   39.3%; Score 398; DB 11; Length 475;
  Best Local Similarity 100.0%; Pred. No. 3.5e-104;
  Matches 398; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   617 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 676
```

```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 60

Qy    677 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGA 736
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGA 120

Qy    737 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 796
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 180

Qy    797 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACA 856
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACA 240

Qy    857 AAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGA 916
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGA 300

Qy    917 GCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 976
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 360

Qy    977 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
          |||||||||||||||||||||||||||||||||||||||
Db    361 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 398


RESULT 5
US-09-910-943-318
; Sequence 318, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
;   APPLICANT: Hemmati-Brivanlou, Ali
;   APPLICANT:  Altman, Curtis
;   TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
;   FILE REFERENCE: 7529/1G148US1
;   CURRENT APPLICATION NUMBER: US/09/910,943
;   CURRENT FILING DATE:  2001-07-23
;   NUMBER OF SEQ ID NOS: 742
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 318
;   LENGTH: 690
;   TYPE: DNA
;   ORGANISM: Xenopus laevis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(690)
;   OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-318

   Query Match              28.5%;  Score 288.8;  DB 9;  Length 690;
   Best Local Similarity    80.5%;  Pred. No. 1.4e-72;
   Matches  338;  Conservative   0;  Mismatches   82;  Indels   0;  Gaps   0;

Qy    595 GCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAG 654
          |||||| ||  |||| ||||||||||||| || ||| | ||||||||| ||  | |||||
```

```
Db        69  GCAGAATTTCTAGAGCAAAATTACGACAGAATATTTAATGACTATGAAAAGCTTCTTCAC  128

Qy       655  TCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGAC  714
              |||||||| |||||| || |||||||||||| | ||||||||| || ||||||||||||||
Db       129  TCTGAGAACTATGTGACGAAGAGACAGTCCCTTAAGCTGCTGGGCGAGCTGATCCTGGAC  188

Qy       715  CGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATG  774
              || |||||||||| |||| ||||| || || || ||||||||| || || |||||| |||||
Db       189  CGACACAACTTTTCCATTATGACTAAATACATAAGCAAGCCTGAAAATCTGAAGCTCATG  248

Qy       775  ATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAG  834
              ||||| || || || |||||| || || |||||| ||||||||||| || ||||| ||||||
Db       249  ATGAATCTGCTCCGTGATAAGAGCCCAAACATTCAGTTTGAAGCATTCCATGTGTTTAAG  308

Qy       835  GTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAG  894
              |||||||| || | ||| |||||||||||||| || |||||| |||||||||||||| ||
Db       309  GTGTTTGTAGCAAATCCAAACAAAACACAGCCCATCGTGGATATCCTGTTAAAAAACCAA  368

Qy       895  CCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTC  954
              |||| | || || || |||||||||||| || || || | || ||||| || ||||||
Db       369  ACCAAGTTAATCGACTTCCTGAGCAGCTTTCAGAAGGATCGAACAGATGACGAACAGTTC  428

Qy       955  GCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA  1014
              | |||||||||||||||||||||||| |||||||| |||||||| || || | | || |||
Db       429  ACCGACGAGAAGAACTACTTGATCAAACAGATACGAGACTTAAAAAAGCCCACGCCATGA  488
```

RESULT 6
US-09-867-701-5263
; Sequence 5263, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5263
;   LENGTH: 435
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-5263

   Query Match   24.3%; Score 246.4; DB 10; Length 435;
   Best Local Similarity   77.6%; Pred. No. 1.8e-60;
   Matches 298; Conservative   0; Mismatches   86; Indels   0; Gaps   0;

```
Qy       626  TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT  685
              ||||     || ||||||||| || ||||| || || || |||||||| || || ||||||||
Db        41  TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC  100

Qy       686  TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA  745
```

```
         | ||||||  || ||  ||  ||   | ||  ||   | |||||||||  | ||| ||||||||| ||| |
Db   101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 160

Qy   746 TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
         ||||  ||  ||  |||||||||  |||  |  ||||||||||||  ||  ||  |||||||||  |||||
Db   161 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220

Qy   806 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
         |||||||||||  ||||||||||  ||||||||||||||||  ||||  ||||  ||||  |||||
Db   221 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280

Qy   866 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
         | ||   |  | ||  ||||||   | ||  ||  |||  ||||||||||  |||||  ||  ||||   || |
Db   281 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340

Qy   926 AAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
         | ||  ||  |||||||||  ||||||||||||      ||||||||||||  |||  ||   |||||||||
Db   341 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400

Qy   986 TCCGAGACTTGAAGAAAACGGCCC 1009
         ||||  ||  |||||||||  |  |  ||  |
Db   401 TCCGGGATTTGAAGAGACCCGCTC 424
```

RESULT 7
US-09-867-701-5899
; Sequence 5899, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5899
;   LENGTH: 447
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-5899

  Query Match   24.1%; Score 244.8; DB 10; Length 447;
  Best Local Similarity 77.3%; Pred. No. 5.3e-60;
  Matches 297; Conservative   0; Mismatches 87; Indels   0; Gaps   0;

```
Qy   626 TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
         ||||      || |||||||||  ||  |||||| ||  ||  ||||||||||  ||  || ||||||||
Db    41 TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 100

Qy   686 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
         | |||||| ||  ||  || ||  || ||  |||  || |  |||||||||  | ||  ||||||||| ||  |
Db   101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 160
```

```
Qy      746  TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
             ||||  ||  ||  |||||||||  |||  |  |||||||||||  ||  ||  || ||||||||  |||||
Db      161  TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220

Qy      806  TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
             ||||||||||  ||||||||||  ||||||||||||||||||||||  ||||  ||||  ||||  || ||||
Db      221  TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280

Qy      866  CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
             | ||  |  || ||||||  | ||  ||  ||| |||||||||||  |||||  ||  ||||  || |
Db      281  CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340

Qy      926  AAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
             | |||  ||  |||||||||  |||||||||||      ||||||||||  |||  ||   |||||||||
Db      341  AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400

Qy      986  TCCGAGACTTGAAGAAAACGGCCC 1009
             || |  ||  |||||||||  |  | ||  |
Db      401  TCAGGGATTTGAAGAGACCAGCTC 424


RESULT 8
US-09-867-701-4953
; Sequence 4953, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4953
;   LENGTH: 450
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-4953

  Query Match            24.1%;   Score 244.8;  DB 10;  Length 450;
  Best Local Similarity   77.3%;   Pred. No. 5.3e-60;
  Matches  297;  Conservative   0;  Mismatches  87;  Indels   0;  Gaps   0;

Qy      626  TTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
             ||||      ||  |||||||||  ||  |||||  ||  ||  |||||||||  ||  ||  |||||||||
Db       27  TTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGTGACAAAAAGACAGTCAC 86

Qy      686  TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
             | ||||||  ||  || ||  || ||  | ||  ||  | |||||||||  | ||| ||||||||||  || |
Db       87  TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAATTATGACAAAATACA 146

Qy      746  TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
             ||||  ||  ||  |||||||||  |||  |  ||||||||||  ||  ||  ||  ||||||||  |||||
Db      147  TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 206
```

```
Qy    806 TCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
          ||||||||||| ||||||||| ||||||||||||||||| |||| |||| |||| || ||||
Db    207 TCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 266

Qy    866 CTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCC 925
          | || | || |||||| | || || ||| ||||||||||| ||||| || |||| || |
Db    267 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 326

Qy    926 AAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGA 985
          | || || ||||||||| ||||||||||   |||||||||| ||| || |||||||||
Db    327 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 386

Qy    986 TCCGAGACTTGAAGAAAACGGCCC 1009
          || | || ||||||| | | ||| |
Db    387 TCAGGGATTTGAAGAGACCAGCTC 410


RESULT 9
US-09-910-943-35
; Sequence 35, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
;   APPLICANT: Hemmati-Brivanlou, Ali
;   APPLICANT:  Altman, Curtis
;   TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
;   FILE REFERENCE: 7529/1G148US1
;   CURRENT APPLICATION NUMBER: US/09/910,943
;   CURRENT FILING DATE:  2001-07-23
;   NUMBER OF SEQ ID NOS: 742
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;   LENGTH: 762
;   TYPE: DNA
;   ORGANISM: Xenopus laevis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(762)
;   OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-35

  Query Match            20.8%;  Score 210.8;  DB 9;  Length 762;
  Best Local Similarity  78.6%;  Pred. No. 4.8e-50;
  Matches  287;  Conservative    0;  Mismatches   75;  Indels   3;  Gaps   3;

Qy      1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          |||||| |||||||| |||||| || || ||||| ||||||||| || || |||||| ||||
Db    397 ATGAAGAAAATGCCATTGTTCAGCAAGTCACATAAAAATCCGGCTGAGATTGTTAAAACT 456

Qy     61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
          ||||| ||||| |||||| | |||||| ||| ||||||||| || || ||||| || |||
Db    457 CTGAAGGACAACATGGCCCTGCTGGAAAGGCAGGACAAAAAAACTGAAAAGGCCTCTGAA 516

Qy    121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
          |||||||||||||| || |||||| | ||||| ||| ||||||| ||| || |||||||
Db    517 GAAGTGTCTAAATCTCTTCAAGCTACAAAAGAGATTTTGTGTGGGACAGGGGACAAAGAA 576
```

```
Qy    181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          || |  |||||| | |||||||||||| |||||||||| |||| ||||||| || | ||
Db    577 CCTCAGACAGAGACGGTGGCTCAGCTCGCACAAGAACTGTACAACAGTGGCTTGTTGGTT 636

Qy    241 ACACTGATAGCTGACC-TGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGAT 299
          || | ||||| ||| |||| || |||||| ||||| || || ||||||||| | |||||
Db    637 ACTTTAATAGCCCACCTTGCATCTCATAGATTTTGANGGCAAGAAAGATGTATCTCAGAT 696

Qy    300 ATTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAG 359
          ||| || |||||| ||||||| |||||| ||||||||||||| | ||||||||||||| |
Db    697 ATTCNAC-ACATCCTGAGAAAACAGATTGGCACTCGGAGTNC-CCTGTGGAGTATATCAA 754

Qy    360 TGCTC 364
          | | |
Db    755 TTCCC 759
```

```
   Query Match              19.2%;  Score 195;  DB 10;  Length 387;
   Best Local Similarity   100.0%;  Pred. No. 1.1e-45;
   Matches  195;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       820  TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       387  TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328

Qy       880  CTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 939
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       327  CTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 268

Qy       940  GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       267  GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208

Qy      1000  AAAACGGCCCCTTGA 1014
              |||||||||||||||
Db       207  AAAACGGCCCCTTGA 193
```

RESULT 11
US-09-880-107-481/c
; Sequence 481, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;   APPLICANT: Horne, Darci T.
;   APPLICANT: Vockley, Joseph G.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Gene Logic, Inc.
;   TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;   FILE REFERENCE: 44921-5028-WO
;   CURRENT APPLICATION NUMBER: US/09/880,107
;   CURRENT FILING DATE: 2001-06-14
;   PRIOR APPLICATION NUMBER: US 60/211,379
;   PRIOR FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: US 60/237,054
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 3950
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 481
;   LENGTH: 387
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA234362
US-09-880-107-481

```
   Query Match              19.2%;  Score 195;  DB 10;  Length 387;
   Best Local Similarity   100.0%;  Pred. No. 1.1e-45;
   Matches  195;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       820  TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       387  TTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328

Qy       880  CTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 939
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327   CTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 268

Qy     940   GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267   GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208

Qy    1000   AAAACGGCCCCTTGA 1014
             |||||||||||||||
Db     207   AAAACGGCCCCTTGA 193
```

RESULT 12
US-10-257-826A-118
; Sequence 118, Application US/10257826A
; Publication No. US20030181407A1
; GENERAL INFORMATION:
;   APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
;   APPLICANT: PALIN, Marie-France
;   APPLICANT: POMAR, Candido
;   APPLICANT: GARIEPY, Claude
;   TITLE OF INVENTION: Steatosis-modulating factors and uses
;   TITLE OF INVENTION: thereof
;   FILE REFERENCE: 14654-2US
;   CURRENT APPLICATION NUMBER: US/10/257,826A
;   CURRENT FILING DATE: 2002-10-17
;   PRIOR APPLICATION NUMBER: 60/197936
;   PRIOR FILING DATE: 2000-04-17
;   PRIOR APPLICATION NUMBER: PCT/CA01/00509
;   PRIOR FILING DATE: 2001-04-12
;   NUMBER OF SEQ ID NOS: 305
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 118
;   LENGTH: 722
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificial sequence
;   OTHER INFORMATION: Muscular steatosis
;   OTHER INFORMATION: Porcine
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(722)
;   OTHER INFORMATION: n = A,T,C or G
US-10-257-826A-118

  Query Match   16.7%; Score 169.8; DB 13; Length 722;
  Best Local Similarity   60.1%; Pred. No. 3.1e-38;
  Matches 303; Conservative   0; Mismatches 196; Indels   5; Gaps   4;

```
Qy     347   TGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCC 406
             ||| | ||      ||| |   |  |  |  || | ||| |  |  |||| || ||
Db       8   TGGTGAATNCCTCTGCCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATNTT 67

Qy     407   CACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTG 466
             |  | |||| |  |   ||| || || ||| | ||| ||||  || ||| |   ||| ||||
Db      68   CCGAAATTNCNNTTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTNNACCGCTTG 127
```

```
Qy    467  CCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTC  524
           ||||||||||  |      |    |   || ||| | |||  ||||  |  || || |  |||
Db    128  CCAAAATCATTTTGNGGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATGTN  187

Qy    525  AACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATA  583
           | || ||| | || ||| | || || | || |  |  ||||| |  |||||  | ||||||
Db    188  ANCATTTTNACATATCTTTACATNCNNTTNCCNCATTTTNNGNNTTACTTTCACGACATA  247

Qy    584  AAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGA  643
           | ||| |      || | || | ||| ||||| |||| || |  |  |||| || ||||||
Db    248  TATTGCTCACNGCGCAANTTTTGGAACANCATTATGATANATTTTTCAGTGAATATGATG  307

Qy    644  AATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGC  703
           ||   |||||| |||  | ||||||||| |  | ||||| || | |||||| || || || |
Db    308  AAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGAAC  367

Qy    704  TGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC  763
           | | ||  |  | |  ||||  ||| ||||| |  ||| || ||  |  | | | |  |||
Db    368  TACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNACC  427

Qy    764  TGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCTTT  822
           |  |  | ||||||||||| | |  || |||||||| || | ||| || ||||| | |||
Db    428  T-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCTTN  486

Qy    823  CATGTTTTTAAGGTGTTTGTGGCC  846
           || |||||||| | |   |||    |
Db    487  CACGTTTTTAANGGGGNTGTNNNC  510
```

RESULT 13
US-10-257-826A-119
; Sequence 119, Application US/10257826A
; Publication No. US20030181407A1
; GENERAL INFORMATION:
; APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
; APPLICANT: PALIN, Marie-France
; APPLICANT: POMAR, Candido
; APPLICANT: GARIEPY, Claude
; TITLE OF INVENTION: Steatosis-modulating factors and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 14654-2US
; CURRENT APPLICATION NUMBER: US/10/257,826A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/197936
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: PCT/CA01/00509
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119
;   LENGTH: 700
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Artificial sequence
;   OTHER INFORMATION: Muscular steatosis

US-10-257-826A-119

```
     Query Match          16.4%;   Score 166.6;   DB 13;   Length 700;
     Best Local Similarity    60.1%;   Pred. No. 2.6e-37;
     Matches  304;   Conservative    0;   Mismatches 197;   Indels    5;   Gaps    4;
```

```
Qy      345 TGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGC 404
                | ||| | ||       ||| |     |   |    |   || | ||| |   |   |||| || ||
Db        6 TCTGGTGAATCCCTCTGCCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATN 65

Qy      405 CCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACT 464
              |   | ||| |   |      ||| || || ||| | ||| ||||  || ||| |   ||| ||
Db       66 TTCCGAAATTTCGATTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTCCACCGCT 125

Qy      465 TGCCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTG 522
              ||||||||||||| |     |     | || || | |||  ||||  |  || || |   ||
Db      126 TGCCAAAATCATTTTGNGGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATG 185

Qy      523 TCAACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACA 581
              | | || ||| | ||   ||| | ||| |   || |   | ||     | |||||   | ||||
Db      186 TNANCATTTTNACATATCTTTACATNCNNTTNCCNCATTTTNNGNNTTACTTTCACGACA 245

Qy      582 TAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGA 641
              || | ||  |       || | |   | || |||||  |||| || |     ||||    || |||||
Db      246 TATATTGCTCACNGCGCAANTTTTGGAACANCATTATGATANATTTTTCAGTGAATATGA 305

Qy      642 GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGA 701
              ||  |||||| ||| | ||||||||||  | |   |||||  || |  ||||| || || ||
Db      306 TGAAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGA 365

Qy      702 GCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAA 761
              ||   | ||   |   |   |  | ||||   |||   ||||| |   || |||  || || | |
Db      366 ACTACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNA 425

Qy      762 CCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCT 820
              |||     | | |||||||||||| | |    || |||||||  ||| |||| |||||| | ||
Db      426 CCT-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCT 484

Qy      821 TTCATGTTTTTAAGGTGTTTGTGGCC 846
              | || |||||||| | |   ||| |
Db      485 TNCACGTTTTTAANGGGGNTGTNNNC 510
```

RESULT 14
US-09-770-445-592

```
   Query Match           15.4%;  Score 156;  DB 9;  Length 861;
   Best Local Similarity  55.8%;  Pred. No. 3.4e-34;
   Matches  319;  Conservative    0;  Mismatches  250;  Indels   3;  Gaps    1;

Qy        394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGA 453
              || | |||| | |  | || || |||| | ||| | |||| | ||||| || || ||
Db         12 GGGTTTGAAAACACCGATATGGCGTTACACTATGGTACTATGTTTAGAGAGTGCATCCGT 71

Qy        454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
              ||| |      |||| |||     |  |  || | ||  | |  ||| ||| | |||
Db         72 CATCAGATTGTTGCAAAATATGTTTTGGACTCGGAGCACGTGAAGAAGTTTTTTTACTAC 131

Qy        514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
               | ||| | | || || |||||| | |||| | ||| ||| |||| ||||| | ||
Db        132 ATACAGCTTCCCAATTTCGACATTGCTGCTGATGCTGCTGCAACTTTTAAGGAACTTCTG 191

Qy        574 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 633
              || || || ||        || || || ||  |    | ||| | |||     |||||| |
Db        192 ACAAGGCACAAGTCTACAGTTGCTGAGTTTCTCATTAAGAATGAAGACTGGTTTTTTGCA 251

Qy        634 GACTATGA---GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAG 690
              ||||| |     ||  | ||| | ||     |||||| ||||     | ||  || | |||
Db        252 GACTACAACTCAAAGCTTCTTGAATCAACTAATTATATTACCCGACGGCAAGCTATTAAG 311

Qy        691 CTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGC 750
              || | || ||  | |  |||| |   ||| || | |||||| |||||| ||| | |||
Db        312 TTGTTGGGCGATATATTATTGGATAGGTCAAATTCGGCTGTGATGACGAAGTATGTGAGC 371
```

```
Qy      751 AAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAG 810
             |  ||| ||| |||    |   ||||||| || ||  | ||   |||    |  || |||
Db      372 TCGATGGATAACTTGAGGATTCTGATGAATCTTCTCAGAGAATCAAGCAAGACTATTCAG 431

Qy      811 TTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATT 870
             | |||||| || |||||||||| ||| ||||||| || |  |    |||| |     ||
Db      432 ATAGAAGCTTTCCATGTTTTTCAAGCTGTTTGTAGCGAACCAAAACAAGCCTTCAGACATC 491

Qy      871 GTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAA 930
             |   | || |||| |   |||    |||  ||| ||  |   || ||  ||    || |
Db      492 GCCAACATTCTGGTGGCAAACAGAAACAAGCTTCTGAGATTGTTGGCTGATATCAAGCCG 551

Qy      931 GAAAGGACGGATGATGAGCAGTTCGCTGACGA 962
             || |     ||| || |    |   ||| |||| |
Db      552 GACAAAGAGGACGAGAGGTTTGACGCAGACAA 583


RESULT 15
US-09-923-876-1251
; Sequence 1251, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN
SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 1251
;    LENGTH: 262
;    TYPE: DNA
;    ORGANISM: Zea mays
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020013958A1 700158378H1
;    NAME/KEY: unsure
;    LOCATION: 148
;    OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-1251

   Query Match              7.3%;  Score 74.2;  DB 9;  Length 262;
   Best Local Similarity   55.5%;  Pred. No. 6.4e-11;
   Matches  142; Conservative    0; Mismatches  114; Indels    0; Gaps    0;

Qy      311 TCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTC 370
             | ||||||    |||| |     |      |  |  || |||||||||    ||||  |
Db        7 TGTTGAGACAGAAGGTTGATGAAAGCTATTGTTGCGTCCAGTATATTGAAAATCATTTTG 66
```

```
Qy   371 ATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGA 430
         || |  ||  | | || |     | ||| |   |   | | || ||   |||||| |
Db    67 ATCTTTTGGATTTCCTTGTTGTTTGCTATAAGAACTTGGAAGTCGCGTTGAATTGTGGAA 126

Qy   431 TTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATC 490
         ||| || ||||||||| ||   | ||      |||||||| |||   || |    || |
Db   127 ACATGTTGCGAGAATGCATAANATATCCTACACTTGCAAAATATATATTGGAGTCAAGCA 186

Qy   491 AATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCT 550
         |||       || ||  |||| || || |||||||||  || |||||||||  || |||||
Db   187 GCTTCGAGTTGTTTTTCCAGTATGTTGAATTGTCAAACTTCGATATTGCATCTGATGCTC 246

Qy   551 TTGCTACTTTCAAGGA 566
         |       |||||||||||
Db   247 TGAACACTTTCAAGGA 262


Search completed: January  6, 2004, 05:04:45
Job time : 1400 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004   Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 01:15:17 ; Search time 2583 Seconds
                                             (without alignments)
                                             9541.130 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt..........tgaagaaaacggccccttga 1014

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_estom:*
                 17: em_gss_hum:*
                 18: em_gss_inv:*
                 19: em_gss_pln:*
                 20: em_gss_vrt:*
                 21: em_gss_fun:*
                 22: em_gss_mam:*
                 23: em_gss_mus:*
                 24: em_gss_pro:*
                 25: em_gss_rod:*
                 26: em_gss_phg:*
                 27: em_gss_vrl:*

```
        28:   gb_gss1:*
        29:   gb_gss2:*


   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.

                                   SUMMARIES
                        %
   Result            Query
      No.    Score   Match Length DB    ID                    Description
   -------------------------------------------------------------------------------
         1    860.4   84.9   1552  11   AK076867           AK076867 Mus muscu
         2    860.4   84.9   2245  11   AK030474           AK030474 Mus muscu
         3    860.4   84.9   3039  11   AK053642           AK053642 Mus muscu
         4    858.8   84.7   1377  11   AK076758           AK076758 Mus muscu
         5    844.8   83.3   1449  11   AK013205           AK013205 Mus muscu
         6    770.6   76.0    822   9   AU125107          AU125107 AU125107
         7    750.8   74.0   1201  13   BX393735           BX393735 BX393735
         8    709.4   70.0   1379  11   AK005323           AK005323 Mus muscu
         9    671.8   66.3    784  10   BG218735           BG218735 RST38476
        10    622.8   61.4   1281  11   AK013161           AK013161 Mus muscu
        11     614    60.6    951  13   BU116522           BU116522 603139786
        12    594.2   58.6    982  13   BQ669953           BQ669953 AGENCOURT
        13    585.2   57.7    934  13   BU518807           BU518807 AGENCOURT
        14    579.2   57.1    713  14   CD354831           CD354831 UI-M-GM0-
        15     578    57.0    742   2   HSM073180         Bx483012 Homo sapi
        16     565    55.7    946  14   CA973078           CA973078 AGENCOURT
        17    553.4   54.6    958  13   BU514920           BU514920 AGENCOURT
        18    550.2   54.3    563   9   AA278473          AA278473 zs81h12.r
        19    542.6   53.5    930  14   CA982682           CA982682 AGENCOURT
        20    536.2   52.9    732  14   CD103801           CD103801 AGENCOURT
        21    533.2   52.6   1186  10   BF159587           BF159587 601769084
   C    22    521.4   51.4    721   9   AW242839          AW242839 xn26f05.x
        23    517.8   51.1    985  13   BU914186           BU914186 AGENCOURT
        24    500.4   49.3   1060  13   BQ899617           BQ899617 AGENCOURT
        25    499.6   49.3    817  14   CB559060           CB559060 AGENCOURT
        26    491.8   48.5    893  13   BU367571           BU367571 603787711
        27    486.6   48.0    842  13   BQ940647           BQ940647 AGENCOURT
        28    473.6   46.7    854  14   CB200465           CB200465 AGENCOURT
   C    29     461    45.5    634  12   BQ021348           BQ021348 UI-H-DH1-
        30    460.8   45.4    933  14   CA983082           CA983082 AGENCOURT
        31    457.6   45.1    575  12   BI499153           BI499153 ie27c05.y
        32     452    44.6    863  13   BU905264           BU905264 AGENCOURT
        33    451.4   44.5    927  12   BI655225           BI655225 603284039
        34    451.2   44.5    876  13   BQ880200           BQ880200 AGENCOURT
        35    445.2   43.9    955  13   BU152376           BU152376 AGENCOURT
        36     445    43.9    534   9   AI645170          AI645170 mt82a02.y
        37    443.8   43.8    833   9   AL961440          AL961440 AL961440
   C    38    442.8   43.7    853  14   CB564338           CB564338 AGENCOURT
        39    441.6   43.6    697  12   BI828691           BI828691 603074707
        40    435.8   43.0   1046  14   CB235478           CB235478 AGENCOURT
        41    435.2   42.9    440   9   AA669484          AA669484 af74g07.r
   C    42    434.2   42.8    662  13   BU631151           BU631151 UI-H-FE1-
        43     433    42.7    961  13   BU129770           BU129770 603117261
        44     430    42.4    752  14   CB524885           CB524885 UI-M-FY0-
        45     428    42.2    985  14   CA970822           CA970822 AGENCOURT
```

RESULT 1
AK076867
```
LOCUS         AK076867              1552 bp    mRNA    linear   HTC 07-DEC-2002
DEFINITION    Mus musculus adult male testis cDNA, RIKEN full-length enriched
              library, clone:4930520C08 product:MO25-LIKE PROTEIN homolog [Homo
              sapiens], full insert sequence.
ACCESSION     AK076867
VERSION       AK076867.1  GI:26345723
KEYWORDS      HTC; CAP trapper.
SOURCE        Mus musculus (house mouse)
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1
  AUTHORS     Carninci,P. and Hayashizaki,Y.
  TITLE       High-efficiency full-length cDNA cloning
  JOURNAL     Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE     99279253
   PUBMED     10349636
REFERENCE     2
  AUTHORS     Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
              Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE       Normalization and subtraction of cap-trapper-selected cDNAs to
              prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL     Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE     20499374
   PUBMED     11042159
REFERENCE     3
  AUTHORS     Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
              Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
              Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
              Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
              Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
              Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
              Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE       RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
  JOURNAL     Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE     20530913
   PUBMED     11076861
REFERENCE     4
  AUTHORS     Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
              Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
              Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
              Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
              Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
              Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
              Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
              Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
              Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
              Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
              Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
```

```
              Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
              Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
              Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
              Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
              Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
              Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
              Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
              and Hayashizaki,Y.
  TITLE       Functional annotation of a full-length mouse cDNA collection
  JOURNAL     Nature 409 (6821), 685-690 (2001)
  MEDLINE     21085660
   PUBMED     11217851
REFERENCE   5
  AUTHORS     The FANTOM Consortium and the RIKEN Genome Exploration Research
              Group Phase I & II Team.
  TITLE       Analysis of the mouse transcriptome based on functional annotation
              of 60,770 full-length cDNAs
  JOURNAL     Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1552)
  AUTHORS     Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
              Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
              Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
              Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
              Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
              Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
              Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
              Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
              Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
              Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
              Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A.,
              Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE       Direct Submission
  JOURNAL     Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
              RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
              Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
              URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
              Fax:81-45-503-9216)
  COMMENT     cDNA library was prepared and sequenced in Mouse Genome
              Encyclopedia Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in RIKEN.
              Division of Experimental Animal Research in Riken contributed to
              prepare mouse tissues.
              Please visit our web site for further details.
              URL:http://genome.gsc.riken.go.jp/
              URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..1552
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4930520C08"
                     /db_xref="MGI:1894876"
                     /db_xref="taxon:10090"
                     /clone="4930520C08"
                     /sex="male"
```

                         /tissue_type="testis"
                         /clone_lib="RIKEN full-length enriched mouse cDNA library"
                         /dev_stage="adult"
          CDS   316. .1320
                         /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                         [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                         98.2%ID, 100%length, match=1002)
                         putative"
                         /codon_start=1
                         /protein_id="BAC36513.1"
                         /db_xref="GI:26345724"
                         /db_xref="MGI:1914081"
                         /translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
                         AMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
                         RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
                         QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
                         YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
                         FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
                         "
          polyA_signal   1539. .1544
                         /note="putative"
          polyA_site   1552
                         /note="putative"
BASE COUNT   490 a   320 c   341 g   401 t
ORIGIN

    Query Match   84.9%;   Score 860.4;   DB 11;   Length 1552;
    Best Local Similarity   90.5%;   Pred. No. 5.8e-176;
    Matches 918; Conservative   0; Mismatches   96; Indels   0; Gaps   0;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   307 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 366

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   367 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 426

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            || |||||| ||||| |||||||||||||||||| ||||||||||||| || ||||| || ||
Db   427 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 486

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
            ||||| ||||||||||||||||||||||| || || || ||||||||||| || |||| |||
Db   487 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 546

Qy   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
            ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   547 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 606

Qy   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
            || ||||||||| ||||||||||||| || || ||| ||||||||||| ||||| || |||
Db   607 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 666

Qy   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
            |||||||||| |||||||||||||||| ||||||||| ||||||||||||||||||||||||
Db   667 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 726

```
Qy    421  CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           || |||||||||||||| |||||  |||||||||||||||| ||||||||||||||||||||
Db    727  CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 786

Qy    481  TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           |||||||||| ||||||||||||| ||| |||||  || ||| |||| || |||||||| |||
Db    787  TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 846

Qy    541  TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
           |||||||||||| |||||||| ||||||| |||||||||||||||||| |||||||||||||
Db    847  TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 906

Qy    601  TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
           |||||||||||||||| ||||||||||||||||||||||||||||| ||||  || ||||||
Db    907  TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 966

Qy    661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
           || |||||  || ||||||||| |||||||| |||||||| |||||||||||||||||| |||
Db    967  AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1026

Qy    721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
           || ||   |||| ||||| |||||||||||||||||| |||||||||||||| ||||||||||
Db    1027 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1086

Qy    781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
           || ||||| || ||||||||||||||||||  || ||||||||| |||||  ||||||||||||
Db    1087 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1146

Qy    841  GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
           ||||||||| || |||||||| |||||||| |||||||| |||||||||||||||||||||||
Db    1147 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1206

Qy    901  CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
           ||||||||||||||||||||||||||  || |||||||||| || || ||||||| ||||||
Db    1207 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1266

Qy    961  GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
           |||||||||||| ||||||||||||| |||||||||||||||| | ||||| |||
Db    1267 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1320
```

RESULT 2
AK030474

```
LOCUS       AK030474                2245 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus adult male pituitary gland cDNA, RIKEN full-length
            enriched library, clone:5330416K15 product:MO25-LIKE PROTEIN
            homolog [Homo sapiens], full insert sequence.
ACCESSION   AK030474
VERSION     AK030474.1  GI:26326468
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
```

TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE   2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409 (6821), 685-690 (2001)
MEDLINE   21085660
PUBMED   11217851
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 2245)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,

Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
        Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
        Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
        Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
        Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
        Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
        Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
        Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
        Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
        Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
        Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
  COMMENT   cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..2245
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:5330416K15"
                     /db_xref="taxon:10090"
                     /clone="5330416K15"
                     /sex="male"
                     /tissue_type="pituitary gland"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS   300..1313
                     /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                     [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                     98.2%ID, 100%length, match=1002)
                     putative"
                     /codon_start=1
                     /protein_id="BAC26978.1"
                     /db_xref="GI:26326469"
                     /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                     SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                     NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                     FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                     SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                     FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                     AAP"
BASE COUNT   667 a   480 c   517 g   581 t
ORIGIN

```
Query Match              84.9%;   Score 860.4;   DB 11;   Length 2245;
Best Local Similarity    90.5%;   Pred. No. 6.2e-176;
Matches  918;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;


Qy       1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     300 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 359

Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
           |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db     360 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 419

Qy     121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
           || ||||| ||||| |||||||||||||||| |||||||||||||| || ||||| || ||
Db     420 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 479

Qy     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           |||||  |||||||||||||||||||||| || || || |||||||||| ||  |||| |||
Db     480 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 539

Qy     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           |||||  ||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     540 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 599

Qy     301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           || ||||||||| ||||||||||||| || ||  ||| ||||||||| ||||| || |||
Db     600 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 659

Qy     361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           |||||||||| ||||||||||||||| |||||||| ||||||||||||||||||||||||
Db     660 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 719

Qy     421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           || ||||||||||||||| ||||| |||||||||||||||| ||||||||||||||||||
Db     720 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 779

Qy     481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           ||||||||||| |||||||||||||| |||||| || ||| |||| || ||||||||| |||
Db     780 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 839

Qy     541 TCAGATGCCTTTGCTACTTTCAAGGATTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
           |||||||||| ||||||||| |||||||| ||||||||||||||||| |||||||||||
Db     840 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 899

Qy     601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
           |||||||||||||||| ||||||||||||||||||||||||||||  ||| || |||||||
Db     900 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 959

Qy     661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
           || ||||| || ||||||||| |||||||| |||||| |||||||||||||||||| |||
Db     960 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1019

Qy     721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
           || || |||| ||||| |||||||||||||||||||| ||||||||||||||| ||||||||
Db    1020 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1079

Qy     781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
```

```
       ||  ||||||  ||  ||||||||||||||||||| ||  ||||||||| |||||  ||||||||||||||
Db    1080 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1139

Qy     841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
           |||||||| ||  |||||||||| |||||||||  ||||||||| |||||||||||||||||||||
Db    1140 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1199

Qy     901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
           |||||||||||||||||||||||  ||  ||||||||||| || || |||||||||  ||||||
Db    1200 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1259

Qy     961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
           ||||||||||||| |||||||||||||| |||||||||||||||| | ||||| |||
Db    1260 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1313
```

RESULT 3
AK053642
LOCUS   AK053642   3039 bp   mRNA   linear   HTC 05-DEC-2002
DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
            library, clone:E130116O21 product:MO25-LIKE PROTEIN homolog [Homo
            sapiens], full insert sequence.
ACCESSION   AK053642
VERSION   AK053642.1 GI:26343600
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4

AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
           Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
           Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
           Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
           Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
           Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
           Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
           Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
           Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
           Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
           Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
           Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
           Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
           Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
           Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
           Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
           Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
           Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
           and Hayashizaki,Y.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL Nature 409 (6821), 685-690 (2001)
   MEDLINE 21085660
    PUBMED 11217851
REFERENCE 5
  AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
           Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
           of 60,770 full-length cDNAs
  JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6   (bases 1 to 3039)
  AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
           Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
           Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
           Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
           Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
           Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
           Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
           Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
           Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
           Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
           Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
           Physical and Chemical Research (RIKEN), Laboratory for Genome
           Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
           RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
           Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
           URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
           Fax:81-45-503-9216)
  COMMENT cDNA library was prepared and sequenced in Mouse Genome
           Encyclopedia Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in RIKEN.
           Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues.
           Please visit our web site for further details.
           URL:http://genome.gsc.riken.go.jp/

```
FEATURES             Location/Qualifiers
     source          1. .3039
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:E130116O21"
                     /db_xref="taxon:10090"
                     /clone="E130116O21"
                     /tissue_type="eyeball"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="0 day neonate"
     CDS             287. .1300
                     /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                     [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                     98.2%ID, 100%length, match=1002)
                     putative"
                     /codon_start=1
                     /protein_id="BAC35457.1"
                     /db_xref="GI:26343601"
                     /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                     SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                     NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                     FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                     SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                     FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
                     AAP"
     polyA_signal    3022. .3027
                     /note="putative"
     polyA_site      3039
                     /note="putative"
BASE COUNT       882 a    645 c    676 g    836 t
ORIGIN
```

```
    Query Match              84.9%;  Score 860.4;  DB 11;  Length 3039;
    Best Local Similarity    90.5%;  Pred. No. 6.6e-176;
    Matches  918;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;


Qy       1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     287 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 346


Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
           |||||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db     347 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 406


Qy     121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
           || ||||| ||||| ||||||||||||||||| ||||||||||||| || |||||| || ||
Db     407 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 466


Qy     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           ||||| |||||||||||||||||||||| || || || |||||||||| || |||| |||
Db     467 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 526


Qy     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           ||||| |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     527 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 586
```

```
Qy    301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          || ||||||||| |||||||||||| || || ||| ||||||||| ||||| || |||
Db    587 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 646

Qy    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          ||||||||||| ||||||||||||||| |||||||| ||||||||||||||||||||||||
Db    647 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 706

Qy    421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
          || ||||||||||||| ||||| |||||||||||||| ||||||||||||||||||||||||
Db    707 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 766

Qy    481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          ||||||||||| |||||||||||||| ||||| || ||| |||| || ||||||||| |||
Db    767 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 826

Qy    541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
          ||||||||||| |||||||| |||||||| |||||||||||||||||| |||||||||||||
Db    827 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 886

Qy    601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          |||||||||||||||||| ||||||||||||||||||||||||||||| |||| || ||||||
Db    887 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 946

Qy    661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
          || ||||| || |||||||| |||||||| |||||||| ||||||| ||||||||||||| |||
Db    947 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1006

Qy    721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
          || || |||| ||||| |||||||||||||||| |||||||||||||||| ||||||||||
Db   1007 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1066

Qy    781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
          || ||||| || ||||||||||||||||||||| || ||||||| ||||| |||||||||||
Db   1067 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1126

Qy    841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
          |||||||| || |||||||||| ||||||| |||||||| |||||| ||||||||||||||||
Db   1127 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1186

Qy    901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
          ||||||||||||||||||||||||| || |||||||||||| || || ||||||||| ||||||
Db   1187 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1246

Qy    961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
          |||||||||||| |||||||||||||| ||||||||||||||| | |||||| |||
Db   1247 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1300
```

RESULT 4
AK076758

```
LOCUS       AK076758                1377 bp    mRNA     linear   HTC 07-DEC-2002
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4930433N18 product:MO25-LIKE PROTEIN homolog [Homo
            sapiens], full insert sequence.
ACCESSION   AK076758
```

```
VERSION     AK076758.1  GI:26345637
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
```

REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
             Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6   (bases 1 to 1377)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
             Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
             Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
             Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
             Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
             Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
             Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
             Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
             Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
             Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
             Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A.,
             Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
             Physical and Chemical Research (RIKEN), Laboratory for Genome
             Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
             RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
             Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
             URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
             Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site for further details.
             URL:http://genome.gsc.riken.go.jp/
             URL:http://fantom.gsc.riken.go.jp/.
FEATURES   Location/Qualifiers
     source   1..1377
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4930433N18"
                     /db_xref="MGI:1894867"
                     /db_xref="taxon:10090"
                     /clone="4930433N18"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS   287..1300
                     /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                     [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                     98.2%ID, 100%length, match=1002)
                     putative"
                     /codon_start=1
                     /protein_id="BAC36470.1"
                     /db_xref="GI:26345638"
                     /db_xref="MGI:1914081"

/translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
AAP"
BASE COUNT   430 a   294 c   306 g   347 t
ORIGIN

    Query Match   84.7%;   Score 858.8;   DB 11;   Length 1377;
    Best Local Similarity   90.4%;   Pred. No. 1.3e-175;
    Matches 917; Conservative   0; Mismatches   97; Indels   0; Gaps   0;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   287 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 346

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
             |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   347 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 406

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
             || ||||| ||||| |||||||||||||| |||||||||||||| || ||||| || ||
Db   407 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 466

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
             ||||| ||||||||||||||||||||| || || || |||||||||| || |||| |||
Db   467 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 526

Qy   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
             ||||| ||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   527 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 586

Qy   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
             || |||||||| ||||||||||||| || || ||| ||||||||| ||||| || |||
Db   587 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 646

Qy   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
             ||||||||| |||||||||||||||||| ||||||| |||||||| |||||||||||||||
Db   647 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 706

Qy   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
             || |||||||||||||| ||||| |||||||||||||| |||||||||||||||||||||
Db   707 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 766

Qy   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
             |||||||||| ||||||||||||||| ||||| || ||| |||| || |||||||| |||
Db   767 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 826

Qy   541 TCAGATGCCTTTGCTACTTTCAAGGATTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
             |||||||||| ||||||||| |||||||| |||||| ||||||||||| ||||||||||||
Db   827 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 886

Qy   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
             ||||||||||||||||| ||||||||||||||||||||||||||| |||| || ||||||
Db   887 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 946

```
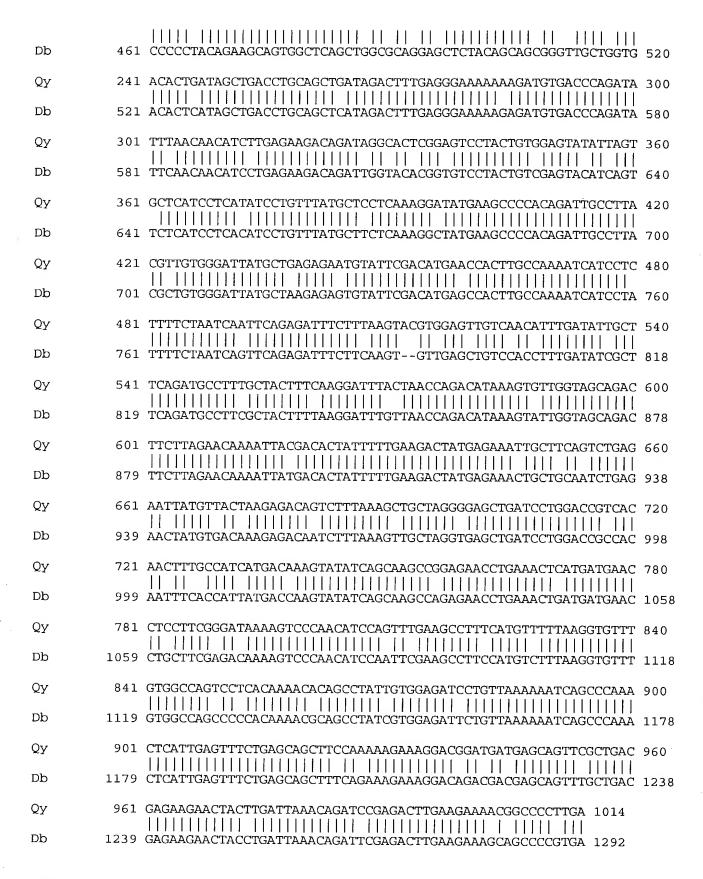Qy      661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
            || |||||| || ||||||||| ||||||||| |||||||| |||||||||||||||| |||
Db      947 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1006

Qy      721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
            || ||   |||| |||||| |||||||||||||||||| |||||||||||||| ||||||||
Db     1007 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1066

Qy      781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
            || |||||| || |||||||||||||||||||| || |||||||| ||||| |||||||||||
Db     1067 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1126

Qy      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
            |||||||| || ||||||||| |||||||| |||||||| ||||||||||||||||||||||
Db     1127 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1186

Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
            |||||||||||| |||||||||| || |||||||||||| || || ||||||||| ||||||
Db     1187 CTCATTGAGTTTTTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1246

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
            |||||||||||| ||||||||||||||| |||||||||||||||||| | ||||| |||
Db     1247 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1300


RESULT 5
AK013205
LOCUS       AK013205                1449 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
            enriched library, clone:2810430N08 product:MO25-LIKE PROTEIN
            homolog [Homo sapiens], full insert sequence.
ACCESSION   AK013205
VERSION     AK013205.1  GI:12850419
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
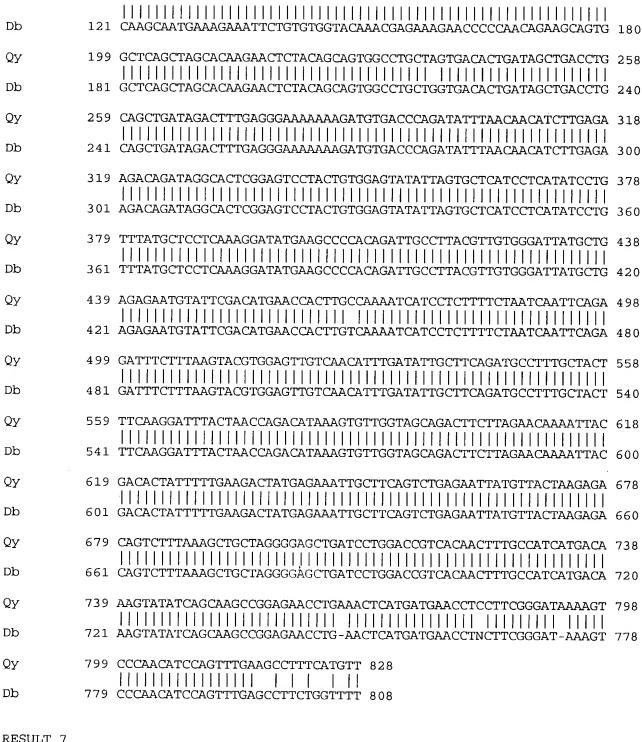```

```
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1449)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
```

COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 7.5 and subtraction to
            Rot = 37.5. Second strand cDNA was prepared with the primer adapter
            of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.
FEATURES   Location/Qualifiers
     source   1. .1449
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2810430N08"
                     /db_xref="MGI:1893512"
                     /db_xref="taxon:10090"
                     /clone="2810430N08"
                     /tissue_type="whole body"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10, 11 days embryo"
     misc_feature   281. .1292
                     /note="MO25-LIKE PROTEIN homolog [Homo sapiens]
                     (SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
                     match=1002)
                     putative"
                     /db_xref="MGI:1914081"
BASE COUNT   453 a   304 c   325 g   367 t
ORIGIN

   Query Match   83.3%;   Score 844.8;   DB 11;   Length 1449;
   Best Local Similarity   90.2%;   Pred. No. 1.4e-172;
   Matches 915; Conservative   0; Mismatches   97; Indels   2; Gaps   1;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
            ||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||||
Db   281 ATGAAAAAAATGCCCTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 340

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
            |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   341 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 400

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
            || |||||| ||||| ||||||||||||||| ||||||||||||| || |||||| || ||
Db   401 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 460

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240

```
                      |||||  |||||||||||||||||||| ||  || ||  |||||||||| ||   ||||  |||
Db      461 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 520

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
                |||||  ||||||||||||||||||| ||||||||||||||||||  ||||||||||||||
Db      521 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAGAGATGTGACCCAGATA 580

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
                || |||||||||  ||||||||||||| ||  ||  |||  |||||||||  ||||| ||  |||
Db      581 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 640

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
                ||||||||||  |||||||||||||  ||||||||   ||||||||||||||||||||||||||||||
Db      641 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 700

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
                ||  ||||||||||||||  |||||  ||||||||||||||  |||||||||||||||||||||||
Db      701 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 760

Qy      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
                ||||||||||  ||||||||||||||||  ||||    ||  |||  ||||  ||  |||||||||  |||
Db      761 TTTTCTAATCAGTTCAGAGATTTCTTCAAGT--GTTGAGCTGTCCACCTTTGATATCGCT 818

Qy      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
                |||||||||||  |||||||||  |||||||||    |||||||||||||||||  ||||||||||||
Db      819 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 878

Qy      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
                |||||||||||||||||| |||||||||||||||||||||||||||||  |||| ||  ||||||
Db      879 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 938

Qy      661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
                ||  ||||| ||  ||||||||||   ||||||||| |||||||   |||||||||||||||||  |||
Db      939 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 998

Qy      721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
                ||  ||   ||||  ||||||  |||||||||||||||||||  ||||||||||||||  ||||||||||
Db      999 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1058

Qy      781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
                ||  |||||  ||  |||||||||||||||||||  ||  |||||||||  ||||| ||||||||||||
Db     1059 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT 1118

Qy      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAA 900
                ||||||||  ||  |||||||| |||||||||  ||||||||  |||||||||||||||||||||||||
Db     1119 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCAAA 1178

Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
                |||||||||||||||||||||||||||||||  |||||||||||| ||  |||||||||| ||||||
Db     1179 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1238

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
                |||||||||||| ||||||||||||||  |||||||||||||||  | ||||| |||
Db     1239 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1292
```

RESULT 6

AU125107

LOCUS	AU125107	822 bp	mRNA linear EST 01-AUG-2002
DEFINITION	AU125107 NT2RM4 Homo sapiens cDNA clone NT2RM4001047 5', mRNA
	sequence.
ACCESSION	AU125107
VERSION	AU125107.1 GI:10949823
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
            Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
            Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
            ,Y., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES            Location/Qualifiers
     source         1. .822
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="NT2RM4001047"
                    /cell_type="teratocarcinoma"
                    /cell_line="NT2"
                    /clone_lib="NT2RM4"
                    /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
                    precursor cells"
BASE COUNT      268 a    164 c    171 g    216 t       3 others
ORIGIN
```

   Query Match   76.0%; Score 770.6; DB 9; Length 822;
   Best Local Similarity   98.5%; Pred. No. 1.5e-156;
   Matches 798; Conservative   0; Mismatches 10; Indels 2; Gaps   2;

```
Qy       19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

Qy       79 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

Qy      139 CAAGCAATGAAAGAAATTCTGTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
```

```
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180

Qy   199 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
         |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   181 GCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

Qy   259 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

Qy   319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

Qy   379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

Qy   439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
         ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480

Qy   499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

Qy   559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600

Qy   619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy   679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy   739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
         ||||||||||||||||||||||||||||| ||||||||||||||||| ||||||||| |||||
Db   721 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778

Qy   799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
         |||||||||||||||||  |  |    | ||
Db   779 CCCAACATCCAGTTTGAGCCTTCTGGTTTT 808


RESULT 7
BX393735
LOCUS       BX393735                1201 bp    mRNA     linear   EST 13-MAY-2003
DEFINITION  BX393735 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC002YI01 5-PRIME, mRNA sequence.
ACCESSION   BX393735
VERSION     BX393735.1  GI:30624044
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 6951.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC002AE01QP1&cluster=6951.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DC002AE01QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DC002YI01"
                     /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT       348 a     223 c     239 g     321 t      70 others
ORIGIN
```

```
   Query Match              74.0%;  Score 750.8;  DB 13;  Length 1201;
   Best Local Similarity    91.0%;  Pred. No. 3.1e-152;
   Matches  766;  Conservative   45;  Mismatches   27;  Indels   4;  Gaps   2;

Qy         1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
             |||||||||||||||||||||||||||| :||||||||||||||||||||||||||||||||
Db       323 ATGAAAAAAATGCCTTTGTTTAGTWAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 382

Qy        61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       383 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 442

Qy       121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
             |:||||||||:|:||||||||  |||:|||::|||  |||:|||||||||||:|||||:||||
Db       443 GWAGTGTCTWAWTCACTGCTAGCWATGWWAGATATTYTGTGTGGTACAWACGAGWAAGAT 502

Qy       181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
             |||||||||||||| |||||||||| |||:|||: :||| |||::|| ||||||||||
Db       503 CCCCCAACAGAAGCTGTGGCTCAGCTTGCAYAAGWTYTCTTCAGYWGTTGCCTGCTAGTG 562

Qy       241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
             ||||||||:|||||||||||||||||||||||||:||||||||::|||||||||||||||
Db       563 ACACTGATWGCTGACCTGCAGCTGATAGACTTTKAGGGAAAARDAGATGTGACCCAGATT 622
```

```
Qy    301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 TTTTACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 682

Qy    361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 742

Qy    421 CGTTGTGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 CGTTGTGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 802

Qy    481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    803 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 862

Qy    541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
          |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db    863 TCAGATGCCTTTGCTACTTTCAAGGATTTWCTAACCAGACATAAAGTGTTGGTAGCAGAC 922

Qy    601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          ||||||||||||||||||||||||||||||||||:::||:||||||||||||||||||||
Db    923 TTCTTAGAACAAAATTACGACACTATTTTTKWWGAYTATGAGAAATTGCTTCAGTCTGAG 982

Qy    661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTA-GGGGAGCTGATCCTGGACCGTCA 719
          ||||||||||||||||||||||||||||||||||||: ||||:|||||||||||||||||
Db    983 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTHGGGGGRGCTGATCCTGGACCGTCA 1042

Qy    720 CAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAA 779
          |||||||||||||||||||||||||:|||::| ::||||| : :||:::::||||:|||||||
Db   1043 CAACTTTGCCATCATGACAAAGTWTATYMYCYMBCCGGSBYHCCYSWWACTCMTGATGAA 1102

Qy    780 CCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTT 839
          || |:| ||| ||: || |:|:||||||||||||||||| : ||||: :|| ||
Db   1103 CCCCYTCGGGGTAAAR---KCCCAMAWCCAGTTTGAAGCCTTTWTKTTTTTWKGKGTTTT 1159

Qy    840 TG 841
          ||
Db   1160 TG 1161
```

RESULT 8
AK005323
LOCUS   AK005323   1379 bp   mRNA   linear   HTC 05-DEC-2002
DEFINITION Mus musculus adult male cerebellum cDNA, RIKEN full-length enriched
            library, clone:1500031K13 product:MO25-LIKE PROTEIN homolog [Homo
            sapiens], full insert sequence.
ACCESSION   AK005323
VERSION   AK005323.1 GI:12837793
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.

TITLE   High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED   10349636
REFERENCE   2
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED   11042159
REFERENCE   3
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED   11076861
REFERENCE   4
AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409 (6821), 685-690 (2001)
MEDLINE   21085660
PUBMED   11217851
REFERENCE   5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6 (bases 1 to 1379)
AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,

```
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
  COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence[5'
            GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA
            was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
            end: SstI. Host: SOLR.
FEATURES             Location/Qualifiers
     source          1..1379
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:1500031K13"
                     /db_xref="MGI:1901050"
                     /db_xref="taxon:10090"
                     /clone="1500031K13"
                     /sex="male"
                     /tissue_type="cerebellum"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             285..1175
                     /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                     [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                     98.2%ID, 100%length, match=1002)
                     putative"
                     /codon_start=1
                     /protein_id="BAB23953.2"
                     /db_xref="GI:26342524"
                     /db_xref="MGI:1916258"
                     /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                     PLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
```

NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
FKNSVFITNRIHGLKRWLSS"
    polyA_signal   1361. .1366
            /note="putative"
    polyA_site   1379
            /note="putative"
BASE COUNT   452 a 284 c 291 g 352 t
ORIGIN

 Query Match   70.0%; Score 709.4; DB 11; Length 1379;
 Best Local Similarity 88.9%; Pred. No. 3e-143;
 Matches 767; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
     ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   285 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 344

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
     |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   345 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 404

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
     || ||||| ||| | ||||||||||||| |||||||||||| || ||||| || ||
Db   405 GAGGTGTCAAAACCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 464

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
     ||||| ||||||||||||||||||| || || || ||||||||||| || |||| |||
Db   465 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 524

Qy   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
     ||||| ||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   525 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 584

Qy   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
     || |||||||| |||||||||||| || || ||| |||||||| ||||| || |||
Db   585 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGT 644

Qy   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
     |||||||||| |||||||||||||||| |||||||| |||||||||||||||||||||||
Db   645 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 704

Qy   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
     || ||||||||||||||| ||||| |||||||||||||| |||||||||||||||||||
Db   705 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 764

Qy   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
     |||||||||| |||||||||||||| ||||| || ||| |||| .|| |||||||| |||
Db   765 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 824

Qy   541 TCAGATGCCTTTGCTACTTTCAAGGATTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
     |||||||||| ||||||||| |||||||| ||||||||||||||||| |||||||||||
Db   825 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 884

Qy   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
     |||||||||||||||||| ||||||||||||||||||||||||||||| |||| || |||||||

```
Db       885  TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 944

Qy       661  AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
              ||  ||||||  ||  |||||||||  |||||||||||  ||||||||  ||||||||||||||||||||  |||
Db       945  AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1004

Qy       721  AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
              || ||   ||||  |||||  |||||||||||||||||||  ||||||||||||||  |||||||||
Db      1005  AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1064

Qy       781  CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
              ||  |||||  ||  ||||||||||||||||||||||  |||||||||  |||||  ||||||    | |
Db      1065  CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGAATTCT 1124

Qy       841  GTGGCCAGTCCTCACAAAACACA 863
              ||    |  |  |  |  |  |  || |||
Db      1125  GTCTTTATTACAAATAGAATACA 1147


RESULT 9
BG218735
LOCUS        BG218735                784 bp     mRNA     linear    EST 21-APR-2001
DEFINITION   RST38476 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION    BG218735
VERSION      BG218735.1  GI:13744756
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
   ORGANISM  Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 784)
   AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
             Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
             , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
             ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
             ,J., Danzig,J. and Ducar,M.
   TITLE     Creation of genome-wide protein expression libraries using random
             activation of gene expression
   JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
   MEDLINE   21227151
    PUBMED   11329013
COMMENT      Contact: Scott J. Cain
             Athersys, Inc.
             3201 Carnegie Ave, Cleveland, OH 44115, USA
             Tel: 216 431 9900
             Fax: 216 361 9596
             Email: scain@athersys.com
             High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="HT1080"
                     /clone_lib="Athersys RAGE Library"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
```

Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT   265 a   151 c   158 g   209 t   1 others
ORIGIN

  Query Match   66.3%; Score 671.8; DB 10; Length 784;
  Best Local Similarity   96.5%; Pred. No. 3.8e-135;
  Matches 718; Conservative   0; Mismatches   23; Indels   3; Gaps   3;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
           ||||||| |||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   32 ATGAAAGAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGTAAATC 91

Qy   61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   92 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 151

Qy   121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   152 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 211

Qy   181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   212 CCCCCAACAGAAGCAGTGGCCCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 271

Qy   241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   272 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 331

Qy   301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   332 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 391

Qy   361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   392 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCCTA 451

Qy   421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   452 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 511

Qy   481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   512 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 571

Qy   541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   572 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 631

Qy   601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATG-AGAAATTGCTTCAGTCTGA 659
           |||||| | ||| | |||||||||||| ||||||||||||||||||||
Db   632 TTCTTAAACAAAATTACGACACTATTTTTTGAAGACTATGAAGAAATTGCTTCAGTCTGA 691

Qy   660 GAATTATGTTAC-TAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTC 718
           ||||||||||||| |||||||||||||||| |||||||||||||||||||||||| | ||

```
Db      692 GAATTATGTTACTTAAGAGACAGTCTTTAGAGCTGCTAGGGGAGCTGATCCTGAAANGTT 751

Qy      719 ACAACTTTGCCATCATGACAAAGT 742
            |||||||||||| ||||||||||||
Db      752 ACAACTTTGCC-TCATGACAAAGT 774
```

RESULT 10
AK013161
LOCUS   AK013161   1281 bp   mRNA   linear   HTC 05-DEC-2002
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
            enriched library, clone:2810425013 product:MO25-LIKE PROTEIN
            homolog [Homo sapiens], full insert sequence.
ACCESSION   AK013161
VERSION   AK013161.1 GI:12850350
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,

Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.

TITLE   Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409 (6821), 685-690 (2001)
MEDLINE   21085660
PUBMED   11217851
REFERENCE   5

AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.

TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6   (bases 1 to 1281)

AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE   Direct Submission
JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 7.5 and subtraction to
            Rot = 37.5. Second strand cDNA was prepared with the primer adapter
            of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved
            with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
            Host: SOLR.

```
FEATURES             Location/Qualifiers
        source       1. .1281
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:2810425013"
                     /db_xref="MGI:1908997"
                     /db_xref="taxon:10090"
                     /clone="2810425013"
                     /tissue_type="whole body"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="10, 11 days embryo"
        misc_feature 289. .1096
                     /note="MO25-LIKE PROTEIN homolog [Homo sapiens]
                     (SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
                     match=1002)
                     putative"
                     /db_xref="MGI:1922871"
BASE COUNT      383 a    275 c    291 g    332 t
ORIGIN
```

Query Match   61.4%; Score 622.8; DB 11;   Length 1281;
Best Local Similarity 82.8%; Pred. No. 1.7e-124;
Matches 763; Conservative   0; Mismatches 97; Indels 62; Gaps   2;

```
Qy       93 AGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAAGCAATGAAAGA 152
            |||  | | || |         ||||||||||||||| ||||| ||||| |||||||||||||| ||
Db      237 AGAAGACAGGATTTCTAAGGCTTCAGAAGAGGTGTCAAAATCTCTGCAAGCAATGAAGGA 296

Qy      153 AATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCTCAGCTAGCACA 212
            |||||||||||| || ||||| || || ||||| |||||||||||||||||||||| || ||
Db      297 AATTCTGTGTGGAACGAACGACAAGGAGCCCCCTACAGAAGCAGTGGCTCAGCTGGCGCA 356

Qy      213 AGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAGCTGATAGACTT 272
            || ||||||||||| || |||| ||||||||| |||||||||||||||||||||| ||||||||
Db      357 GGAGCTCTACAGCAGCGGGTTGCTGGTGACACTCATAGCTGACCTGCAGCTCATAGACTT 416

Qy      273 TGAGGGAAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGACAGATAGGCAC 332
            ||||||||||||||||||||||||||||||||| ||||||||| |||||||||||| || |||
Db      417 TGAGGGAAAAAAAGATGTGACCCAGATATTCAACAACATCCTGAGAAGACAGATTGGTAC 476

Qy      333 TCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAA 392
            ||| |||||||||| |||| || ||| ||||||||||| ||||||||||||| |||||
Db      477 ACGGTGTCCTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAA 536

Qy      393 AGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCG 452
            ||| ||||||||||||||||||||||||||||| |||||||||||| ||||| ||||||||||
Db      537 AGGCTATGAAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTCG 596

Qy      453 ACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTA 512
            ||||||| ||||||||||||||||||||| ||||||||| ||||||||||||||| |||||
Db      597 ACATGAGCCACTTGCCAAAATCATCCTATTTTCTAATCAGTTCAGAGATTTCTTCAAGTA 656

Qy      513 CGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACT 572
            || |||| ||||| || ||||||||| |||||||||||||| ||||||||| ||||||| |||
Db      657 TGTTGAGCTGTCCACCTTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAG-------- 708
```

```
Qy      573 AACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGA 632
                                                                  ||||||||
Db      709 --------------------------------------------------------ATTTTTGA 716

Qy      633 AGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCT 692
            |||||||||||| |||| || ||||||||| ||||| || ||||||| |||||||| |
Db      717 AGACTATGAGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTT 776

Qy      693 GCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAA 752
            ||||||  ||||||||||||||||||| ||||| || |||| ||||| ||| ||||||||
Db      777 GCTAGGTGAGCTGATCCTGGACCGCCACAATTTCACCATTATGACCAAG--TATCAGCAA 834

Qy      753 GCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTT 812
            |||  ||||||||||||| ||||||||||||| ||||| || |||||||||||||||| ||
Db      835 GCCAGAGAACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATT 894

Qy      813 TGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGT 872
             |||||||| ||||| ||||||||||||||||||||||| || ||||||| |||||||| ||
Db      895 CGAAGCCTTCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGT 954

Qy      873 GGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGA 932
            ||||| ||||||||||||||||||||||||||||||||||||||||||||| || |||||
Db      955 GGAGATTCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTTCAGAAAGA 1014

Qy      933 AAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGA 992
            |||||| || || ||||||||| ||||||||||||||||||| ||||||||||||| |||||
Db     1015 AAGGACAGACGACGAGCAGTTTGCTGACGAGAAGAACTACCTGATTAAACAGATTCGAGA 1074

Qy      993 CTTGAAGAAAACGGCCCCTTGA 1014
            |||||||||| | ||||| |||
Db     1075 CTTGAAGAAAGCAGCCCCGTGA 1096
```

RESULT 11
BU116522
LOCUS   BU116522   951 bp   mRNA   linear   EST 25-NOV-2002
DEFINITION   603139786F1 CSEQCHL15 Gallus gallus cDNA clone ChEST129122 5', mRNA
             sequence.
ACCESSION   BU116522
VERSION   BU116522.1 GI:25323402
KEYWORDS   EST.
SOURCE   Gallus gallus (chicken)
  ORGANISM   Gallus gallus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
             Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 951)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
             Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE   A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT   Contact: Simon Hubbard
             Department of Biomolecular Sciences
             University of Manchester Institute of Science and Technology (UMIST
)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES   Location/Qualifiers
    source   1. .951
                  /organism="Gallus gallus"
                  /mol_type="mRNA"
                  /strain="Compton Line 15I"
                  /db_xref="taxon:9031"
                  /clone="ChEST129122"
                  /sex="Female"
                  /tissue_type="cerebrum"
                  /dev_stage="adult"
                  /lab_host="DH10B"
                  /clone_lib="CSEQCHL15"
                  /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                  EcoRI; Site_2: Not1; Modification of pBluescript II KS(+)
                  [Stratagene] vector to accommodate cDNA produced with the
                  T-trimmed protocol (Construction of uni-directionally
                  cloned cDNA libraries from messenger RNA for improved 3'
                  end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387
                  ,624). Cut pBluescript II KS(+) with NotI and EcoRI.
                  Ligate in double stranded adaptor containing BsgI and
                  BamHI sites [5'ggccgcgtgcagccccggatccgaaaaaaag]
                  [5'aattcttttttttcggatccggggctgcacgc]"
BASE COUNT   303 a   206 c   202 g   240 t
ORIGIN

  Query Match   60.6%; Score 614; DB 13; Length 951;
  Best Local Similarity 82.8%; Pred. No. 1.3e-122;
  Matches 737; Conservative   0; Mismatches 150; Indels   3; Gaps   3;

```
Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
                ||||||||| |||||| |||||||||||||| |||||||||||| || |  || || ||
Db        6 GAAGTGTCAAAATCTCTGCAAGCAATGAAGTAAATTCTGTGTGGGACCACAGACAAGGAG 65

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
                || || |||||||| ||||||||||||| |||||||| | |||| |||||||||| ||||||
Db       66 CCACCGACAGAAGTAGTGGCTCAGCTGGCACAAGAATTGTACAACAGTGGCCTTCTAGTG 125

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
                ||||| || ||  |||||||||| ||||| |||||||| ||||| |||||  ||||||||||
Db      126 ACACTTATTGCCAACCTGCAGCTCATAGATTTTGAGGGTAAAAAGGATGTTTCCCAGATA 185

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
                |||||||||||| ||||||||||| || ||||| || || |||||||||| || ||||||
Db      186 TTTAACAACATCCTGAGAAGACAAATTGGCACACGAAGCCCTACTGTGGAATACATTAGT 245

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
                || |||||| |||||||||||| ||||| || ||||| |||||| |||||| | ||||||||||
Db      246 GCCCATCCACATATCCTGTTCATGCTTCTGAAAGGCTATGAATCCCCAAATATTGCCTTA 305

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAA-AATCATCCT 479
                || ||||| |||||||||||| || || || |||||||||||| | ||||| ||||||| ||
Db      306 CGCTGTGGAATTATGCTGAGGGAGTGCATCCGACATGAACCATTGGCCAACAATCATACT 365
```

```
Qy    480 CTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGC 539
          |||||  |  || |||||||| ||||| ||||| |||||  ||||||||||||||| ||
Db    366 TTTTTCAGAACAGTTCAGAGACTTCTTCAAGTATGTGGAAATGTCAACATTTGATATAGC 425

Qy    540 TTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGA 599
          ||  ||||||||||||||| |||||||| ||  |||| || || ||  ||||||||||||
Db    426 ATCTGATGCCTTTGCTACATTCAAGGACTTGTTAACAAGGCACAAGTTGTTGGTAGCAGA 485

Qy    600 CTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGA 659
          ||  |  ||||||||||||| || || || ||||| || ||||| ||| | ||||| |||||
Db    486 TTTTATGGAACAAAATTATGATACGATCTTTGAGGATTATGAAAAACTCCTTCATTCTGA 545

Qy    660 GAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCA 719
          ||||||  ||  || |||||||||||| |||||||||  ||  |||||  || || ||| | ||
Db    546 GAATTACGTAACAAAGAGACAGTCTTTGAAGCTGCTGGGTGAATTGATTCTAGACAGACA 605

Qy    720 CAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAA 779
          |||||  |||||||||||||| |||||||||||  || ||||| ||||| || ||||||||
Db    606 CAACTTCGCCATCATGACAAAATATATCAGCAAACCAGAGAATCTGAAGCTGATGATGAA 665

Qy    780 CCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTT 839
          | | || ||  ||  || ||||| ||||||||| || |||||||| || |||||  ||||| ||
Db    666 CTTGCTGCGAGACAAAAGCCCCAACATTCAATTTGAAGCATTCCATGTGTTCAAGGTTTT 725

Qy    840 TGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 899
          |||||||||||||  |||||||| ||||| ||||| || |||||||||||| | ||||| ||| ||||
Db    726 TGTGGCCAGTCCAAACAAAACTCAGCCCATCGTGGAGATCCTGCTGAAAAACCAG-CCAA 784

Qy    900 ACTCATTGAGTTTCTGAGCAGCTTCCA-AAAAGAAAGGACGGATGATGAGCAGTTCGCTG 958
          ||||| |||||||||||||  ||||| ||| || ||||||| ||| ||||||||||| | |
Db    785 GCTCATCGAGTTTCTGAGCCATTTCCAGAAACGAGAGGACGGTTGACGAGCAGTTCACCG 844

Qy    959 ACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCC 1008
          ||||||||||||||| |||| || || ||||||||||||||||||  | | |
Db    845 ACGAGAAGAACTACCTGATCAAGCAAATCCGAGACTTGAAGAAGGCCGAC 894
```

RESULT 12
BQ669953

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2407 row: m column: 13
High quality sequence stop: 508.

FEATURES   Location/Qualifiers
     source   1. .982
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6255924"
                      /tissue_type="epidermoid carcinoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_102"
                      /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
                      Site_2: EcoRI; cDNA made by oligo-dT priming.
                      Directionally cloned into EcoRI/XhoI sites using the
                      following 5' adaptor: GGCACGAG(G). Library constructed
                      by Ling Hong in the laboratory of Gerald M. Rubin
                      (University of California, Berkeley) using ZAP-cDNA
                      synthesis kit (Stratagene) and Superscript II RT (Life
                      Technologies). Note: this is a NIH_MGC Library."
BASE COUNT   217 a   200 c   357 g   197 t   11 others
ORIGIN

  Query Match   58.6%;   Score 594.2; DB 13;   Length 982;
  Best Local Similarity 96.5%;   Pred. No. 2.6e-118;
  Matches 628; Conservative   0; Mismatches   20; Indels   3; Gaps   2;

Qy   334 CGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAA 393
          ||||||||||||||||||| ||||||||||||||||||||||||||| |||||||||||||
Db   1 CGGAGTCCTACTGTGGAG-ATATTAGTGCTCATCCTCATATCCTGGTTATGCTCCTCAAA 59

Qy   394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGA 453
          |||||||||||||||||||||||||||||| ||| |||||||||||||||||| ||||||
Db   60 GGATATGAAGCCCCACAGATTGCCTTACATTGGGGGATTATGCTGAGAGAATGGATTCGA 119

Qy   454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 179

Qy   514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 239

Qy   574 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 633
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   240 ACCAGACATAAAGTGGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 299

Qy   634 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTG 693
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   300 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGGTACTAAGAGACAGTCTTTAAAGCTG 359

Qy   694 CTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAG 753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db      360 CTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAG 419

Qy      754 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTT 813
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      420 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTT 479

Qy      814 GAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 873
            |||||||||||| |||||||| | ||||||||||||||||||||||||||||||||||||
Db      480 GAAGCCTTTCATGGTTTTAAGGGGGTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 539

Qy      874 GAGATCCTGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAA 933
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      540 GAGATCCTGGTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAA 599

Qy      934 AGG--ACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAAC 982
            |||    ||||||||||| || |||||||||| ||||||| | |||
Db      600 AGGGACGGGATGATGAGCANNTCCCTGACGAGAAGACTACTTGGGTTAAC 650
```

RESULT 13
BU518807

LOCUS   BU518807   934 bp   mRNA   linear   EST 12-SEP-2002
DEFINITION AGENCOURT_10171930 NIH_MGC_134 Mus musculus cDNA clone
            IMAGE:6516567 5', mRNA sequence.
ACCESSION   BU518807
VERSION   BU518807.1 GI:22826333
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 934)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14095 row: e column: 16
            High quality sequence stop: 656.
FEATURES   Location/Qualifiers
     source   1..934
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6516567"
                     /tissue_type="undifferentiated limb"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_134"
                     /note="Vector: pCMV-SPORT6.1.ccdb; Site_1: EcoRV; Site_2:
                     NotI; Cloned unidirectionally. Primer: Oligo dT. Average
```

insert size 1.7 kb. Constructed by ResGen, Invitrogen
Corp. Note: this is a NIH_MGC Library."
BASE COUNT      301 a      198 c      200 g      234 t        1 others
ORIGIN

    Query Match                57.7%;    Score 585.2;   DB 13;   Length 934;
    Best Local Similarity      89.0%;    Pred. No. 2.3e-116;
    Matches  654;  Conservative      0;  Mismatches   79;   Indels    2;   Gaps    2;

```
Qy        2 TGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCC 61
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db      113 TGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATTC 172

Qy       62 TGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAG 121
            ||||||||||    |||||||||||||||||||||||||||||||||||||||||||||||
Db      173 TGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAG 232

Qy      122 AAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAAC 181
            | ||||| |||||  |||||||||||||||| ||||||||||||| || ||||| || || |
Db      233 AGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAGC 292

Qy      182 CCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTGA 241
            ||||  |||||||||||||||||||||||  ||  || |||||||||| ||   |||| ||||
Db      293 CCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTGA 352

Qy      242 CACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATAT 301
            ||||  ||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      353 CACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATAT 412

Qy      302 TTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTG 361
            | |||||||| ||||||||||||| ||  || | |||||||||||| ||||||  || |||
Db      413 TCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCCTACTGTCGAGTACATCAGTT 472

Qy      362 CTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTAC 421
            ||||||||||| ||||||||||||  |||||||| ||||||||||||||||||||||||||
Db      473 CTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTAC 532

Qy      422 GTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCT 481
            | |||||||||||||| |||||| ||||||||||||||| ||||||||||||||||||| |
Db      533 GCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTAT 592

Qy      482 TTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTT 541
            |||||||||| ||||||||||||||| ||||| || ||| |||| || |||||||||| ||
Db      593 TTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCTT 652

Qy      542 CAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACT 601
            ||||||||| |||||||||| |||||||| ||||| ||||||||||| |||||||||||||
Db      653 CAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGACT 712

Qy      602 TCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGA 661
            |||||||||||||||| |||||||||||||||||||||||||||| ||||   ||||||||
Db      713 TCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGNCATCTGAGA 772

Qy      662 ATTATGTTACTAAGAG-ACAGTCTTTAAAGCTGCTAGGG-GAGCTGATCCTGGACCGTCA 719
            | |||| || |||||| ||| ||||||||| ||||||||| |||||||||| |  |  |
Db      773 ACTATGTGACAAAGAGAACATTCTTTAAAGTTGCTAGGGTGAGCTGATCCCTGGACCGCC 832
```

```
Qy        720 CAACTTTGCCATCAT 734
              ||   ||  || |||
Db        833 CACAATTTTCACCAT 847
```

RESULT 14
CD354831
```
LOCUS          CD354831                  713 bp    mRNA    linear   EST 29-MAY-2003
DEFINITION     UI-M-GM0-cge-i-10-0-UI.r1 NIH_BMAP_GM0 Mus musculus cDNA clone
               IMAGE: 30361641 5', mRNA sequence.
ACCESSION      CD354831
VERSION        CD354831.1  GI:31147332
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
  ORGANISM     Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1  (bases 1 to 713)
  AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
  TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL      Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Dr. Jim Lin, University of Iowa
                cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Distribution information can be found at
               http://genome.uiowa.edu/distribution/mousefl.html
                This clone was contributed by the Brain Molecular Anatomy Project
               (BMAP)
               Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1..713
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE: 30361641"
                     /tissue_type="whole brain"
                     /dev_stage="1, 5 and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GM0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site.Double strand cDNA was size
                     selected according to mRNA size fraction,ligated with EcoR
                     I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
```

BASE COUNT   220 a   159 c   140 g   192 t   2 others
ORIGIN

   Query Match   57.1%;   Score 579.2;   DB 14;   Length 713;
   Best Local Similarity   89.9%;   Pred. No. 4.4e-115;
   Matches 642; Conservative   0; Mismatches   70; Indels   2; Gaps   2;

```
Qy    282 AAAAGATGTGACCCAGATATTTAACAACAT-CTTGAGAAGACAGATAGGCACTCGGAGTC 340
          |||||||||||||||||||||| |||||||| | ||||||||||||||| || || ||| |||
Db      1 AAAAGATGTGACCCAGATATTCAACAACATCCNTGAGAAGACAGATTGGTACACGGTGTC 60

Qy    341 CTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATG 400
          ||||||| ||||| || ||| ||||||||||| |||||||||||||| ||||||||| ||||
Db     61 CTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATG 120

Qy    401 AAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAAC 460
          |||||||||||||||||||||| ||||||||||||||||| ||||| ||||||||||||| |
Db    121 AAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGC 180

Qy    461 CACTTGCCAAAATCATCCTCTTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGT 520
          |||||||||||||||||||| |||||||||| ||| |||||||||||||| ||||| || |||
Db    181 CACTTGCCAAAATCATCCTATTTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGC 240

Qy    521 TGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGAC 580
          |||| || |||||||| |||||||||||||| ||||||||| |||||||| ||||||||| |||||||||
Db    241 TGTCCACCTTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGAC 300

Qy    581 ATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATG 640
          ||||||| ||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    301 ATAAAGTATTGGTAGCAGACTTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATG 360

Qy    641 AGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGG 700
          ||||| |||| || ||||||||| ||||| || ||||||||| ||||||||| ||||||| |
Db    361 AGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTG 420

Qy    701 AGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGA 760
          ||||||||||||||| ||||| || |||| |||| |||||||||||||||||||||| ||||
Db    421 AGCTGATCCTGGACCGCCACAATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGA 480

Qy    761 ACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCT 820
          |||||||||| |||||||||||| ||||| || |||||||||||||||||||| || |||||||
Db    481 ACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCT 540

Qy    821 TTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCC 880
          | ||||| ||||||||||||||||||||||| || |||||||| ||||||| ||||||||| |
Db    541 TCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTC 600

Qy    881 TGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGG 940
          ||||||||||||||||||||||||||||||||||||||||||||| || |||||||||||| |
Db    601 TGTTAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAG 660

Qy    941 ATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACT 994
          | || ||||||||| ||||||||||||| |||||| |||||| ||||||| |||||||
Db    661 ACGACGAGCAGTTTGCTGACGAGAAG-ACTACCTGATTANACAGATTCGAGACT 713
```

RESULT 15
HSM073180
ID   HSM073180   standard; RNA; EST; 742 BP.
XX
AC   BX483012;
XX
SV   BX483012.1
XX
DT   09-MAY-2003 (Rel. 75, Created)
DT   09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX
DE   Homo sapiens mRNA; EST DKFZp686C08234_r1   (from clone DKFZp686C08234)
XX
KW   EST; expressed sequence tag.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-742
RA   Ottenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.W., Weil B.,
RA   Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RT   ;
RL   Submitted (07-MAY-2003) to the EMBL/GenBank/DDBJ databases.
RL   MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
XX
CC   This is the 5' sequence of the clone insert
CC   Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC   Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC   sequenced by MediGenomix (Martinsried/Germany) within the cDNA
CC   sequencing consortium of the German Genome Project.
CC   No s1 sequence available.
CC   This clone (DKFZp686C08234) is available at the RZPD in Berlin.
CC   Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC   14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
FH   Key   Location/Qualifiers
FH
FT   source   1. .742
FT   /db_xref="taxon:9606"
FT   /mol_type="mRNA"
FT   /organism="Homo sapiens"
FT   /clone="DKFZp686C08234"
FT   /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT   DH10B; sites SfiIA + SfiIB"
FT   /dev_stage="adult"
FT   /tissue_type="cDNA-collection"
XX
SQ   Sequence 742 BP; 256 A; 143 C; 162 G; 179 T; 2 other;

   Query Match   57.0%; Score 578; DB 2; Length 742;
   Best Local Similarity   99.8%; Pred. No. 8e-115;
   Matches 578; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db     164 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 223

Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     224 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 283

Qy     121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     284 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 343

Qy     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     344 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACTCTACAGCAGTGGCCTGCTAGTG 403

Qy     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     404 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 463

Qy     301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     464 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 523

Qy     361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     524 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 583

Qy     421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     584 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 643

Qy     481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
           |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     644 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGNTGTCAACATTTGATATTGCT 703

Qy     541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGA 579
           |||||||||||||||||||||||||||||||||||||||
Db     704 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGA 742
```

Search completed: January 6, 2004, 03:18:09
Job time : 2589 secs